

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**



Replacement Sheet

Figure 1

ID-65

Clone 3-60

5

(SEQ ID NO: 1)

10

15

20

25

30

35

40

GTGTTTATGATGAAAAAAGGACAAGTAAATGATACTAAGCAA
TCTTACTCTCTACGTAAATATAAATTTGGTTTAGCATCAGTAA
TTTTAGGGTCATTTCATAATGGTCACAAGTCCTGTTTTTGCGGA
TCAAACACTACATCGGTTCAAGTTAATAATCAGACAGGCACTAG
TGTGGATGCTAATAATTCTTCCAATGAGACAAGTGCGTCAAGT
GTGATTACTTCCAATAATGATAGTGTTCAAGCGTCTGATAAAG
TTGTAAATAGTCAAAATACGGCAACAAAGGACATTACTACTC
CTTTAGTAGAGACAAAGCCAATGGTGGAAAAAACATTACCTG
AACAAGGGAATTATGTTTATAGCAAAGAAACCGAGGTGAAAA
ATACACCTTCAAAATCAGCCCCAGTAGCTTTCTATGCAAAGA
AAGGTGATAAAGTTTTCTATGACCAAGTATTTAATAAAGATA
ATGTGAAATGGATTTTCATATAAGTCTTTTGGTGGCGTACGTCG
ATACGCAGCTATTGAGTCACTAGATCCATCAGGAGGTTCAGA
GACTAAAGCACCTACTCCTGTAACAAATTCAGGAAGCAATAA
TCAAGAGAAAATAGCAACGCAAGGAAATTATACATTTTTCACA
TAAAGTAGAAGTAAAAAATGAAGCTAAGGTAGCGAGTCCAA
CTCAATTTACATTGGACAAAGGAGACAGAATTTTTTTACGACC
AAATACTAACTATTGAAGGAAATCAGTGGTTATCTTATAAATC
ATTCAATGGTGTTCGTCGTTTTGTTTTGCTAGGTAAAGCATCTT
CAGTAGAAAAAACTGAAGATAAAGAAAAAAGTGTCTCCTCAAC
CACAAGCCCGTATTACTAAAACCTGGTAGACTGACTATTTCTAA
CGAAACAACACTACAGGTTTTGATATTTTAATTACGAATATTAAA
GATGATAACGGTATCGCTGCTGTTAAGGTACCGGTTTGGACTG
AACAAGGAGGGCAAGATGATATTAAATGGTATACAGCTGTAA
CTACTGGGGATGGCAACTACAAAGTAGCTGTATCATTTGCTG
ACCATAAGAATGAGAAGGGTCTTTATAATATTCATTTATACTA
CCAAGAAGCTAGTGGGACACTTGTAGGTGTAAACAGGAACCTAA
AGTGACAGTAGCTGGAACCTAATTCTTCTCAAGAACCTATTGA
AAATGGTTTACCAAAGACTGGTGTTTATAATATTATCGGAAGT
ACTGAAGTAAAAAATGAAGCTAAAATATCAAGTCAGACCCAA
TTTACTTTAGAAAAAGGTGACAAAATAAATTATGATCAAGTA
TTGACAGCAGATGGTTACCAGTGGATTTCTTACAAATCTTATA
GTGGTGTTCGTCGCTATATTCCTGTGAAAAAGCTAACTACAAG
TAGTGAAAAAGCGAAAGATGAGGCGACTAAACCGACTAGTTA
TCCCAACTTACCTAAAACAGGTACCTATACATTTACTAAAACCT
GTAGATGTGAAAAAGTCAACCTAAAGTATCAAGTCCAGTGGA
TTTAATTTTCAAAAGGGTGAAAAAATACATTATGATCAAGTGT

TAGTAGTAGATGGTCATCAGTGGATTTCATACAAGAGTTATTC
CGGTATTCGTCGCTATATTGAAATTTAA

5 (SEQ ID NO: 2)
MFMMKKGQVNDTKQSYSLRKYKFGLASVILGSFIMVTSPVFADQTTSVQVNN
QTGTSVDANNSSNETSASSVITSNNDSVQASDKVVNSQNTATKDITPLVETK
PMVEKTLPEQGNVYVYSKETEVKNTPSKSAPVAFYAKKGDKVFYDQVFNKDN
VKWISYKSFGGVRRYAAIESLDPSGGSETKAPTPVTNSGSNNQEKIATQGNYT
10 FSHKVEVKNEAKVASPTQFTLDKGDRIFYDQILTIEGNQWLSYKSFNGVRRFV
LLGKASSVEKTEDKEKVSPQPQARITKTGRLTISNETTTGFDILITNIKDDNGIA
AVKVPVWTEQGGQDDIKWYTAVTTGDGNYKVAVVSFADHKNEKGLYNIHLY
YQEASGTLVGVTGKVTVAGTNSSQEPIENGLPKTGVYNIIGSTEVKNEAKISS
QTQFTLEKGDKINYDQVLTADGYQWISYKSYSGVRRYIPVKKLTTSSEKAKD
15 EATKPTSYPNLPKTGTYTFTKTVDVKSQPKVSSPVEFNFQKGEKIHVDQVLVV
DGHQWISYKSYSGIRRYIEI*

Sequence description

20 A) Length: 1642 bp - 547 aa (full length gene)
B) Sequence Characteristics:
Potential leader peptide sequence
Orf is preceded by a potential Shine-
Dalgarno sequence.

25 ID-66

Clone 3-5

30 (SEQ ID NO: 3)
ATGATATTGAGACGTCGAACTATTGTTTTATGGCAACTGGGTATCGCCATT
TCTCTCATTCTTAGTATTCTAGCCTTAAATCTTTATTTCCATAGTACTCCCTT
GCAAACCAATGCAGCTTTACGGAACCTTGCTCCTTCATTAAACCATCTTTT
TGGGACAGATGGTTTAGGTAGGGATATGTTTGTGACGACGATTAAAGGAC
35 TTTATTTCTCTCTACAAGTCGGCTTATTAGGTGCCCTTATGGGGGGTCATTCT
GGCGACAGTTTTTTGGAGTGCTTGCAGGTTTAGGAAATAGCATTATTGATAA
AATAATAGCATGGTTAGTTGATTTGTTTATTGGTATGCCTCATTTGATTTTT
ATGATTCTCATTTCTTTTGTGTTGGGAAAGGTGCTCAAGGGGGTCATCATT
GCAACGGCTGTTACACATTGGCCTTCTTTAGCAAGGCTTATCCGCAATGAA
40 GTCTATCATCTAAAGAATAAAGAATTTGTCCAACCTTTCTAAAAGTATGGGA
AAAACGCCTTATTATATTGTGAGGCATCATATCCTGCCTTTGATTGCTTCTC
AAATTTTCATTGGTTTTATCCTCTTATTTCCACATGTCATCCTACATGAAGC
ATCAATGACTTTCTTAGGATTTGGGCTCTCTGCCGAACAACCTTCGGTTGG
TATCATTCTGTCAGAGGCAGCTAAGCATATCTCTCTTGGAATTGGTGGTT

GGTTATCTTTCCAGGACTTTATCTTATTTTGGTTGTCAATGCATTTGATACT
ATCGGAGAATCTTTAAAGAAACTCTTTTACCCTCAAACCTGATCATTTTTAG

5 (SEQ ID NO: 4)
MILRRRTIVLWQLGIAISLILSILALNLYFHSTPLQTNAALRNLAAPSLNHLFGTD
GLGRDMFVRTIKGLYFSLQVGLLGALMGVILATVFGVLAGLGNSIIDKILAWL
VDLFIGMPHLIFMILISFVVGKGAQGVIIATAVTHWPSLARLIRNEVYHLKNKE
FVQLSKSMGKTPYYIVRHHILPLIASQIFIGFILLFPHVILHEASMTFLGFGLSAE
10 QPSVGIILSEAAKHISLGNWWLVIFPGLYLILVVNAFDTIGESLKKLFYPQTDHF
*

Sequence description

15 A) Length: 822 bp - 274 aa (full length gene)
B) Sequence Characteristics:
Potential leader peptide sequence
Orf is preceded by a potential Shine-
20 Dalgarno sequence.

ID-78

25 Clone 3-5b

(SEQ ID NO: 5)
ATGACAGAAACATTATTAAGCATTAAAGACCTCTCCATCACCTTCACTCAA
TACGGAAGATTTTTTAAAACCATTTCAATCAACACCGATACAAGCGCTGA
30 ATTTAGAAATTAAAAAAGGTGAGTTATTAGCTATTATAGGTGCTAGTGGTT
CGGGGAAGAGTTTATTAGCACATGCTATTATGGATATTCTTCCTAAAAATG
CATCTGTAACAGGAGATATGATTTATCGTGGTCAATCACTAAATTCTAAAC
GCATTAAACAGTTGCGAGGAAAAGATATTACGTTGATTCCACAATCAGTT
AATTATTTAGATCCATCTATGAAAGTCAAACATCAGGTGCGCTTAGGTATC
35 TCAGAAAATTCAAAGGCTACTCAAGAAGGATTGTTTCAACAGTTTGGTTTA
AAAGAAAGTGATGGTGACTTGGATCCTTTCCAACCTTTCTGGCGGAATGCTC
CGACGTGTTTTGTTTACAACGTGTATTAGTGATAAGGTTTCTTTGATTATTG
CGGATGAGCCCACCCCTGGATTACATCCAGATGCTCTGCAAATGGTTTTAG
ACCAACTACGCTCCTTTGCAGATAAAGGAATAAGCGTTATATTTATCACTC
40 ATGATATTGTAGCAGCTAGTCAAATTGCTGATCGTATTACTATTTTTAAAG
AGGGAAAAGCTATTGAAACAGCTCCAGCTAGTTTCTTTAGCGGAAATGGA
GAGCAGTTACAAACAGAATTTGCTAGAAGTTTATGGCGCTCTCTCCACAG
CAAGAATTTTTGAAAGGAGTTACTCATGACCTTAGAGGCTAA

(SEQ ID NO: 6)

MTETLLSIKDL SITFTQYGRFLKPFQSTPIQALNLEIKKGELLAIGASGSGKSL
AHAIMDILPKNASVTGDMIYRGQSLNSKRIKQLRGKDITLIPQSVNYLDPSMK
VKHQVRLGISENSKATQEGLFQQFGLKESDGDLDPFQLSGGMLRRVLFTTCIS
5 DKVSLIIADEPTPGLHPDALQMVL DQLRSFADKGISVIFITHDIVAASQIADRITI
FKEGKA IETAPASFFSGNGEQLQTEFARSLWRS LPPQQEFLKGVTHDLRG*

10 Sequence description

A) Length: 804 bp - 268 aa (full length gene)

B) Sequence Characteristics:

No obvious leader peptide sequence

15 Orf is preceded by a potential Shine-
Dalgarno sequence.

This gene was not isolated using the LEEP
system. However in determining a full length
gene sequence for ID-76, this gene was
20 identified downstream and fully sequenced.

ID-79

25 Clone 3-5c

(SEQ ID NO: 7)

GTCCATCTGGGGTGGTTCCCGATTGGTATTTCTTCTCCGATAGGTACTTTGA
GTCAAGATATTACGTTAGCTGATCGTATTAAGCACCTTATTTTACCTGTTTT
30 CACGGTAAGTATTCTAGGCATTGCCAATGTA ACTCTTCATACTAGAACTAA
AATGATGTCGGTACTTTCTAGTGAATATGTCTTATTTGCCAGAGCGCGTGG
GGAAACGGAATGGCAAATTTTTTAAAAATCATTGTCTTAGAAATGCTATCGT
ACCAGCTATTACACTGCATTTTTTCCTATTTTGGAGAATTGTTTGGAGGATCC
GTTCTTGCTGAGCAAGTTTTCTCATATCCAGGACTAGGGTCTACCCTAACT
35 GAAGCAGGACTTAAAAGTGATACACCGCTACTTCTAGCTATTGTGATGATA
GGGACATTATTTGTTTTTGC GGGCAATCTTATTGCGGATATTTTAAATAGC
ATAATCAATCCACAGTTAAGGAGAAAAGTATGA

(SEQ ID NO: 8)

40 VHLGWFPIGISSPIGTLSDITLADRIKHLILPVFTVSI LGIANVTLHTRTKMMSV
LSSEYVLFARARGETEWQIFKNHCLRNAIVPAITLHFSYFGELFGGSVLA EQVF
SYPGLGSTLTEAGLKSDTPLLLAIVMIGTLFVFAGNLIADILNSIINPQLRRKV*

Sequence description

A) Length: 495 bp - 165 aa (partial gene sequence)

B) Sequence Characteristics:

N-terminus has yet to be determined.

5 This gene was not isolated using the LEEP system. However in determining a full length gene sequence for ID-76, this gene was identified upstream.

10

ID-80

Clone 2-17

15

(SEQ ID NO: 9)

TTGCGGACAATTACGTTCAAACACAATGAAACGCGATCGTCAAAAAGCGA
AGGTAGGGCGGTAATGCTTAAAAGATTATTTACTGAAGATGGGGAATTGA
CAAAGATTAGTCGTCGTTTCGTTTGGATGTTAGTGGTTATCTATTGTCTTAT
TATTGTCAGGATGTGTTTTGGGCCTCAAATTATGATTGAGGGGGGTATCAAC
20 TCCGAATGTTTACGCGCTTCGGAAGAATTGTAGCTCTTTTAGTACCATTTAA
TTCTTTTCGTAGTTTAGATCAGCTAACTAGCTTTAAAGAGATTTTTTGGGTT
ATTGGTCAAATGTAGTGAATATTTTACTGCTGTTTCCTCTCATTATAGGGT
TACTATCCCTAAAGCCAAGTTTACGGAAATATAAAAGCGTTATATTACTTG
CTTTCTTGATGTCTCTTTTCATAGAGTGTACTCAAGTTGTTTTAGATATTTT
25 AATAGATGCTAATCGGGTTTTTGAAATCGACGATCTATGGACAAATACCTT
AGGCGGTCCTTTCGCCCTATGGAGTTATCGAAACATAAAAGGTTGGCTTCT
AACTATTAGAAAATGA

30

(SEQ ID NO: 10)

MRTITFKHNETRSSKSEGRAVMLKRLFTEDGELTKISRRFVWMLVVIYCLIIVR
MCFGPQIMIEGVSTPNVQRFGRIVALLVPFNSFRSLDQLTSFKEIFWVIGQNVV
NILLFPLIIGLLSLKPSLRKYKSVILLAFLMSLFIECTQVVLDILIDANRVFEIDD
LWTNTLGPFALWSYRNIKGWLLTIRK*

35

Sequence description

A) Length: 579 bp - 193 aa (full length gene)

B) Sequence Characteristics:

40

Possesses a potential leader peptide sequence

No obvious Shine-Dalgarno, but the 'TTG' codon may not be the actual translation start point.

A methionine (ATG) that occurs ~22 codons downstream of the 'TTG' is preceded by a

potential Shine-Dalgarno sequence and may
represent the actual start codon.

ID 81

5

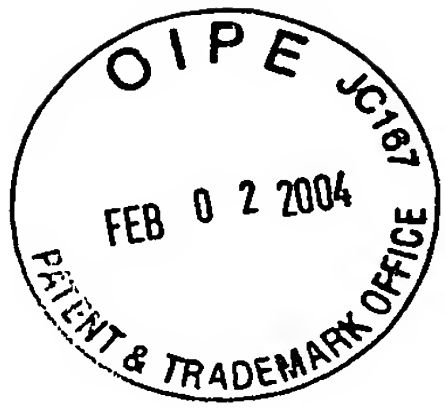
Clone 3-1

(SEQ ID NO: 11)

10 TTGAAAAATTTAAATCGTTATGTAGTTGCGGTTTCTGGAGTCGTTTTACATT
TAATGCTAGGATCAACTTATGCTTGGAGTGTGTTTCGTAACCCAATTATCT
CAGAGACTGGTTGGGATATTTTCATCAGTTTCATTCGCTTTTAGTTTGGCTAT
TTTTTGTCTAGGAATGTCTGCAGCTTTTATGGGACACTTAGTAGAGCGTTTT
GGTCCTAGGATAATGGGAATGATTTCTGCTATTTTATATGGAGCAGGGAAT
GTGTTAACAGGCTTAGCCATTGAAACTCAGCAGTTATGGTTACTGTATGTT
15 GCATACGGTATTTTAGGAGGAATCGGACTTGGTTCAGGTTATATTACTCCA
GTATCGACTATTATTAAATGGTTTCCTGATAGGAGGGGACTAGCAACAGG
ATTCGCTATTATGGGATTTGGCTTTGCTTCTTTAGTAACAAGTCCGCTTGCA
CAATCCTTACTGATTAGGATTGGTGTGGGTAAAACGTTTTTATATTTTGGGA
TTAGTATATTTTTTTTGTTCATGATGATTGCCTCACAATTTATTAAACAACCAC
20 CTCAGGAAAAAATAACTATTTTGACTCACGATGGTAAAAAGAATGCTATG
AATTCACAAATTATCACTGGATTAAAAGCAAACGTCGCTATAAAATCAAA
AACCTTTTACATCATTTGGTTGACCTTGTTTATTAATATTTTCGTGTGGCTTA
GGTTTAATATCAGCAGCTTCACCAATGGCACAAGATTTAGCAGGCTATTCC
GCAGAATCTGCAGCCTTATTAGTAGGGGTACTAGGGGATTTTAACGGTTTTT
25 GGACGTCTGTTATGGGCAAGTCTCTCTGACTACATTGGACGCCCCGTTGACC
TTTATAATATTATTTATTGTGAACCTTATTATGACTTCTAGTTTATTTTTGTC
ATTCAATGCTATTGTATTTGCAATAGCGATGTCTATTTTAATGACTTGTTAT
GGTGCAGGTTTTTCCTTATTACCTGCTTATCTAAGTGATATTTTTGGAACAA
AGGAATTAGCTACTTTACATGGTTATAGTTTAACAGCATGGGCAATAGCAG
30 GTCTGTTTGGGCCCCTATTGTTATCAAAGACATATTCATGGGGAAATTCCT
ATCAATTGACATTAATGGTTTTTTGGTTTTTTTATTCTTATTCGGATTATTGTTA
TCTCTATATTTAAGAAAATTAACAACCTAAAGTTGTGTAG

(SEQ ID NO: 12)

35 LKNLNRVYVAVSGVVLHLMLGSTYAWSVFRNPIISETGWDISSVSFAFSLAIFC
LGMSAAFMGHLVERFGPRIMGMISAILYGAGNVLTGLAIETQQLWLLYVAYG
ILGGIGLGSYITPVSTIIKWFPDRRGLATGFAIMGFGFASLVTSPLAQSLLIRIG
VGKTFYILGLVYFFVMMIASQFIKQPPQEKITLTHDGKKNAMNSQIITGLKAN
VAIKSKTFYIIFWLTLFINISCGLGLISAASPMAQDLAGYSAESAALLVGVLGIFN
40 GFGRLLWASLSDYIGRPLTFIILFIVNFIMTSSLFLSFNAIVFAIAMSILMTCYGA
GFSLLPAYLSDIFGTKELATLHGYSLTAWAIAAGLFGPLLLSKTYSWGNSYQLTL
MVFGFLFLFGLLLSLYLRKLTTKVV*



Sequence description:

5

A] Length 1221 bp - 407 a.a (full length gene).

B] TTG start codon with Shine-Dalgarno sequence upstream. Obvious signal peptide, with hydropathy plot exhibiting many possible membrane spanning regions, indicating protein to be transmembrane.

10

ID-82

Clone 48

15

(SEQ ID NO: 13)

ATGGCAGATAAAAAACAGAACATTTAAACTTGTAGGTGCAGGATCTTCTAG
CACACAAGAAAAAATTGAAAAGCCTGCTCTTTCGTTTATGCAAGATGCGTG
GCGTCGCTTGAAAAAAAAACAAATTAGCAGTAGTTTCACTCTATTTATTAGC
TCTTTTACTTACTTTTTCGTTAGCCTCAAATTTATTTGTAACCTCAGAAGGAT
GCTAATGGGTTTGATTCGAAAAAAGTAACGACATATCGCAACTTACCACCT
AAATTGAGTTCAAACCTTCCTTTTTTGAATGGTAGCATTAAATCCATCA

20

(SEQ ID NO: 14)

MADKNRTFKLVGAGSSSTQEKIEKPALSFMQDAWRRLKKNKLAVVSLYLLA
LLTFSLASNLFTQKDANGFDSKKVTTYRNLPPKLSSNLPFWNGSINPS

25

Sequence description:

30

A] Current length is 303 bp - 101 aa

B] No obvious signal peptide but Shine

Dalgarno sequence upstream of the ATG start

codon. Not identified directly using the LEEP system but was found directly downstream of ID-34 described in WO 00/06736.

35

ID-83

Clone 98

40

(SEQ ID NO: 15)

ATGAAAATAGTAGTACCAGTAATGCCTCGCAGTCTTGAAGAGGCTCAAGA
AATAGATTTATCAAAATTTGATAGTGTTGATATTATTGAATGGCGAGCTGA

5 TGCCTTACCAAAGGATGACATTATTAATGTAGCTCCAGCTATTTTTGAGAA
ATTCGCAGGTCATGAAATTATTTTTACTTTTCGTACAACGCGTGAAGGTGG
TAATATTGTCTTATCTGATGCTGAGTATGTTGAGTTAATCCAGAAAATTAA
TTCTATCTACAATCCAGATTATATTGATTTTGAGTATTTTTCACATAAAGAA
GTTTTTCAAGAAATGCTAGAATTTCCAAATTTAGTCCTGTCTTATCACAATT
TTCAAGAGACACCGGAGAATATTATGGAGATATTTTCAGAATTAACAGCC
CTAGCACCACGAGTTGTGAAAATCGCAGTAATGCCAAAGAATGAACAAGA
TGTCTTAGACGTTATGAATTACACTCGCGGTTTCAAGACTATTAATCCTGA
TCAAGTTTATGCGACGGTATCTATGAGTAAAATTGGACGTATTTCTCGTTTT
10 GCTGGTGATGTA ACTGGATCTAGTTGGACATTTGCATATTTAGATTCATCT
ATCGCACCCGGACAAATTACTATTTTCAGAGATGAAGCGTGTCAAAGCATT
GCTTGACGCTGACTGA

(SEQ ID NO: 16)
15 MKIVVPVMPRSLEEAEIDLSKFDSVDIIEWRADALPKDDIINVAPAIFEFK
HEIIFTFRTTREGGNIVLSDAEYVELIQKINSIYNPDYIDFEYFSHKEVFQEMLEF
PNLVLSYHNFQETPENIMEIFSELTALAPRVVKIAVMPKNEQDVLDVMNYTRG
FKTINPDQVYATVSMISKIGRISRFAGDVTGSSWTFAYLDSSIAPGQITISEMKRV
KALLDAD*
20

Sequence description:

25 A] Length 678 bp, 225 aa (full length gene)
B] No obvious signal peptide, but there is a
Shine Dalgarno immediately upstream of ORF.

30 ID-84

Clone RS-52

(SEQ ID NO: 17)
35 ATGAAAGACTTATTTGCAACAACAGAAGCATCATCAAGGAAACAGGAACA
AGATAGAATTGTCAATTACATAAAACAACATGTTGAGTTAACAAATGGTA
ATCAAATAAAAAAAAAAATTGAGTTTATCGACTTTCAAAAAAAAAATGAGATGACA
GGTACATGGGGAATTTCTACTAAAATTAATGAACAATTTTCGATTAGTTTT
TCTGAAGATAGAATTGGTGGTAACTTAGAGCATTAGGATATCAACCGAA
TGAAATAGGTTTTTTCAAAGGACATCAATAGTAATAATCAAAATGTTAATGA
40 TATTGAAGTGATTTATATGAAGAAAGAATAG

(SEQ ID NO: 18)

MKDLFATTEASSRKQEQDRIVNYIKQHVELTNGNQIKKIEFIDFQKNEMTGTW
GISTKINEQFSISFSEDRIGGKLRALGYQPNEIGFSKDINSNNQNVNDIEVIYMK
KE*

5

Sequence description:

A] length: 333 bp - 111 aa (partial sequence)

B] No obvious Shine Dalgarno sequence upstream
of the ATG start codon, and no obvious signal
peptide within the protein.

10

ID-85

15

Clone RS-53

(SEQ ID NO: 19)

ATGAAAAAACGTATATGGTATTTGATAATAATAATCACAGTAATTTTAGGA
GGACTAGCCATGAAAAACTTATTTGCAACAACAGAAGCATCATCAAGGAA
ACAGGAACAAGATAGAATTGTCAATTACATAAAACAACATGTTGAGTTAA
CAAATGGTAATCAAATAAAAAAAATTGAGTTTATCGACTTTCAAAAAAAT
GAGATGACAGGTACATGGGGAATTTCTACTAAAATTAATGAACAATTTTCG
ATTAGTTTTTCTGAAGATAGAATTGGTGGTAAACTTAGAGCATTAGGATAT
CAACCGAATGAAATAGGTTTTTCAAAGGACATCAATAGTAATAATCA

20
25

(SEQ ID NO: 20)

MKKRIWYLIITVILGGLAMKNLFATTEASSRKQEQDRIVNYIKQHVELTNGN
QIKKIEFIDFQKNEMTGTWGISTKINEQFSISFSEDRIGGKLRALGYQPNEIGFSK
DINSNNQ

30

Sequence description:

A] Length: 351 bp - 117 aa (Partial sequence)

B] Obvious signal peptide and Shine Dalgarno
sequence upstream of the ATG start codon.

35

40

ID-86

Clone ID-74

(SEQ ID NO: 21)

ATGTCAAATCAATATGATTATATCGTTATTGGTGGAGGTAGT
GCAGGCAGTGGTACCGCTAATAGGGCAGCCATGTATGGAGC
AAAAGTCCTGTTAATTGAAGGTGGACAAGTAGGTGGAACCTG
5 TGTAACTTAGGTTGTGTACCTAAGAAAATCATGTGGTATGG
TGCACAAGTTTCTGAGACACTCCATAAGTATAGTTCAGGTTA
TGGTTTTGAAGCCAATAATCTTAGTTTTGATTTTACTACTCTA
AAAGCTAATCGCGATGCTTACGTGCAGCGGTCTAGACAGTCG
TATGCCGCTAATTTTGAAGCGTAATGGGGTTCGAAAAGATTGAT
10 GGATTTGCTCGTTTTATTGATAACCATACTATTGAAGTGAATG
GTCAGCAATATAAAGCTCCTCACATTACTATTGCAACAGGTG
GACACCCTCTTTACCCTGATATTATTGGAAGTGAACCTTGGTG
AGACTTCTGATGATTTTTTTTGGATGGGAGACCTTACCAAATTC
TATATTGATTGTTGGGGCGGGCTATATCGCGGCAGAACTTGC
15 TGGAGTGGTTAATGAATTAGGCGTTGAAACCCATCTTGCATT
TAGAAAAGACCATATTCTACGCGGATTTGATGACATGGTAAC
AAGTGAGGTTATGGCTGAAATGGAGAAATCAGGTATCTCTTT
ACATGCTAACCATGTACCTAAATCTCTTAAACGCGATGAAGG
TGGCAAGTTGATTTTTTGAAGCTGAAAATGGGAAAACGCTTGT
20 CGTTGATCGTGTAATATGGGCTATCGGCCGTGGACCAAATGT
AGACATGGGACTTGAAAATACCGATATTGTTTTAAATGATAA
AGATTATATCAAAACAGATGAATTTGAGAATACTTCTGTAGA
TGGCGTGTATGCTATTGGAGATGTTAATGGGAAAATTGCCTT
GACACCGGTAGCAATTGCAGCAGGTCGTCGCTTATCAGAAAG
25 ACTTTTAAATCATAAAGATAACGAAAAATTAGATTACCATAA
TGTACCTTCAGTTATTTTTACTCACCTGTAAATGGGACGGTA
GGACTTTCAGAAGCAGCAGCTATCGAGCAATTTGGAAAAGAT
AATATCAAAGTCTATACATCAACTTTTACCTCTATGTATACGG
CTGTTACCAGTAATCGCCAAGCAGTTAAGATGAAGCTCATAA
30 CCCTAGGAAAAGAGGAAAAAAGTTATTGGGCTTCATGGTGTTG
GTTATGGTATTGATGAAATGATTCAAGGTTTTTCAGTTGCTAT
CAAAATGGGGGCTACTAAAGCAGACTTTGATGATACTGTTGC
TATTCACCCAACCTGGATCTGAGGAATTTGTTACAATGCGCTA
A

35

(SEQ ID NO: 22)

MSNQYDYIVIGGGSAGSGTANRAAMYGAKVLLIEGGQVGGTC
VNLGCVPKKIMWYGAQVSETLHKYSSGYGFEANNLSFDFTTLK
ANRDAYVQRSRQSYAANFERNGVEKIDGFARFIDNHTIEVNGQ
40 QYKAPHITIAATGGHPLYPDIIGSELGETSDDFFGWETLPNSILIVG
AGYIAAELAGVVNELGVETHLAFRKDHILRGFDDMVTSEVMAE
MEKSGISLHANHVPKSLKRDEGGKLFEAENGKTLVVDRVIWAI
GRGPNVDMGLENTDIVLNDKDYIKTDEFENTSVDGVYAIGDVN
GKIALTPVAIAAGRRLSERLFNHKDNEKLDYHNVPSVIFTHPVIG

TVGLSEAAAIEQFGKDNIKVYTSTFTSMYTA VTSNRQAVKMKLI
TLGKEEKVIGLHGVGYGIDEMIQGF SVAIKMGATKADFDDTVAI
HPTGSEEFVTMR*

5

ID-87

Clone RS-55

10

(SEQ ID NO: 23)

15

20

25

30

35

40

ATGACAAAAAACATCTTAAAACGCTTGCCTTGGCACTTACTACAGTATCA
GTAGTGACATACAGCCAGGAGGTATATGGATTAGAAAGAGAGGAATCGGT
CAAACAAGAACAAACCCAGTCAGCTTCAGAAGATGATTGGTTCGAAGAAG
ATAATGAGAGGAAAACAAATGTTTCTAAAGAGAATTCTACTGTTGATGAA
ACAGTTAGTGATTTATTTTCTGATGGAAATAGTAATAACTCTAGTTCTAAA
ACCGAGTCAGTGGTAAGTGACCCTAAACAAGTCCCCAAAGCAAAACCAGA
GGTTACACAAGAAGCAAGCAATTCTAGTAATGATGCTAGCAAAGTAGAAG
TACCAAAACAGGATACAGCTTCAAAAAAGGAAACTCTAGAAACATCAACT
TGGGAGGCCAAAAGATTTCGTAAGTACTAGAGGGGATACTTTAGTAGGTTTTCA
AAATCTGGAATTAATAAGTTATCTCAAACATCACACTTGGTTTTTACCAAGT
CATGCAGCAGATGGAAGTCAATTGACACAAGTAGCTAGCTTTGCTTTTACT
CCAGATAAAAAGACGGCCATTGCAGAATATACAAGTAGGCTAGGAGAAA
ATGGGAAACCGAGTCGTTTAGATATTGATCAGAAGGAAATTATTGATGAG
GGAGAAATATTTAATGCTTACCAGTTGACTAAGCTTACTATTCCAAATGGT
TATAAGTCTATTGGTCAAGATGCTTTTGTGGACAATAAGAATATTGCTGAG
GTTAACCTTCCTGAGAGTCTCGAGACTATTTTCAGACTATGCTTTTGCTCACA
TGTCTTTAAAACAAGTAAAGTTACCAGATAACCTAAAGGTCATTGGAGAA
TTAGCTTTTTTTTGATAATCAGATTGGTGGTAAGCTTTACTTGCCACGTCACT
TGATAAAATTAGCAGAACGCGCTTTCAAATCTAATCGTATTCAAACAGTTG
AATTTTTGGGAAGTAAGCTTAAGGTTATAGGAGAAGCAAGTTTTCAAGAT
AATAATCTGAGGAATGTTATGCTTCCGGATGGACTTGAAAAAATAGAATC
AGAAGCTTTTACAGGAAATCCAGGAGATGAACATTACAACAATCAGGTTG
TATTGCGCACAAGGACAGGCCAAAATCCACATCAACTTGCGACTGAGAAT
ACTTACGTCAATCCGGACAAATCATTGTGGCGTGCAACACCTGATATGGAT
TATACCAAATGGTTAGAGGAAGATTTTACCTATCAAAAAAATAGTGTTACA
GGTTTTTCAAATAAAGGCTTACAAAAGGTAAGACGTAATAAAAACTTAGA
AATTCCAAAACAACACAATGGTATTACTATTACTGAAATTGGTGATAACGC
TTTTCGCAATGTTGATTTTCAAAGTAAACTTTACGTAAATATGATTTGGA
AGAAATAAAGCTCCCCCTCAACTATTCGGAAAATAGGTGCTTTTGCTTTTCA
ATCTAATAACTTGAAATCCTTTGAAGCAAGTGAAGATTTAGAAGAGATTA
AAGAGGGAGCCTTTATGAATAATCGTATTGGAAGTCTAGACTTGAAAGAC
AACTTATCAAATAGGTGATGCTGCTTTCCATATTAATCATATTTATGCC
ATTGTTCTTCCAGAATCTGTACAAGAAATAGGACGTTTCAGCTTTTCGACAA
AATGGTGCGCTTCACCTTATGTTTATCGGAAATAAGGTAAAACAATTGGT

GAAATGGCTTTTTTATCCAATAAACTGGAAAGTGTAATCTCTCTGAGCAA
AAACAATTAAAGACAATTGAGGTCCAAGCTTTTTTCGGATAATGCCCTTAGT
GAAGTAGTCTTACCGCCAAATTTACAGACTATTCGTGAAGAGGCTTTCAAA
AGGAATCATTTGAAAGAAGTGAAGGGTTCATCTACATTATCTCAGATTACT
5 TTTAATGCTTTTGATCAAAATGATGGGGACAAACGCTTTGGTAAGAAAGTG
GTTGTTAGGACACATAATAATTCTCATATGTTAGCAGATGGTGAGCGTTTT
ATCATTGATCCAGATAAGCTATCTTCTACAATGGTAGACCTTGAAAAGGTT
TTAAAAATAATCGAAGGTTTAGATTACTCTACATTACGTCAGACTACTCAA
ACTCAGTTTAGAGAAATGACTACTGCAGGTAAAGCGTTGTTATCAAAATCT
10 AACCTCCGACAAGGAGAAAAACAAAAATTCCTTCAAGAAGCACAATTTTT
CCTTGGTCGCGTTGATTTGGATAAAGCCATAGCTAAAGCTGAGAAGGCTTT
AGTGACCAAGAAGGCAACAAAGAATGGTCATTTGCTTGAGAGGAGTATTA
ACAAAGCGGTATTAGCTTATAATAATAGTGCTATTAATAAGCTAATGTTA
AGCGCTTGGAAGAGAGTTAGACTTGCTGACAGATTTAGTCGAGGGGAAAA
15 GGACCATAGCGCAAGCTACAATGGTACAAGGAGTTTATTTATTAAAGAC
GCCTTTACCATTGCCAGAATATTATATCGGATTGAACGTTTATTTTGACAA
GTCTGGAATAATTGATTTATGCACTTGATATGAGTGATACTATTGGCGAGGG
ACAAAAAGATGCATATGGTAATCCTATATTAAATGTTGACGAGGATAATG
AAGGTTATCATACCTTGGCAGTTGCCACTTTAGCTGATTATGAAGGTCTTT
20 ATATTAAAGATATTTTAAATAGTTCCCTTGATAAGATTAAAGCAATACGCC
AGATTCCTTTGGCAAAATATCATAGATTAGGAATTTTCCAAGCTATCCGAA
ATGCAGCGGCAGAAGCAGACCGATTGCTTCCTAAGACACCTAAGGGGTAC
CTAAATGAAGTCCCAAATTATCGTAAAAACAAATGGAGAAAAATTTAAA
ACCAGTTGATTATAAAACGCCGATTTTAAATAAGGCTTTACCTAATGAAAA
25 GGTAGACGGTGATAGAGCGGCTAAAGGTCATAATATAAATGCGGAGACTA
ATAATTCTGTAGCTGTAAACACCAATAAGGTCCGAGCAGCAATTACATAAGT
CACAGTCTGATGTAAATTTACCTCAAACAAGTTCTAAAAATAATTTTATAT
ACGAGATTCTAGGATACGTTAGTTTATGTTTGCTTTTCCTAGTAAGTGTGG
GAAAAAAGGAAAACGAGCAAGAAAATAA

30 (SEQ ID NO: 24)
MTKKHLKTLALALTTVSVVTYSQEVYGLEREESVKQEQTQSASEDDWFEEDN
ERKTNVSKENSTVDETVDLFDSDGNSNNSSSKTESVVSDPKQVPKAKPEVTQE
ASNSSNDASKVEVPKQDTASKKETLETSTWEAKDFVTRGDTLVGFSKSGINKL
35 SQTSHLVLP SHAADGTQLTQVASFAFTPDKKTAIAEYTSRLGENGKPSRLDIDQ
KEIIDEGEIFNAYQLTKLTIPNGYKSIGQDAFVDNKNIAEVNLPESLETISDYAF
AHMSLKQVKLPDNLKVIGELAFFDNQIGGKLYLPRHLIKLAERAFKSNRIQTV
EFLGSKLKVIGEASFQDNNLRNVMLPDGLEKIESEAFTGNPGDEHYNNQVVLR
TRTGQNP HQ LATENTYVNPDKSLWRATPDMDYTKWLEEDFTYQKNSVTGFS
40 NKGLQKVR RNKNLEIPKQHNGITITEIGDNAFRNVDFQSKTLRKYDLEEIKLPS
TIRKIGAF AFQSNNLKSFEASEDLEEIKEGAFMNNRIGTLDLKD KLIKIGDAAFH
INHIYAIVLPESVQEIGRSAFRQNGALHLMFIGNKVK TIGEMAFLSNKLESVNL
SEQKQLKTIEVQAFSDNALSEVVLPPNLQTIREEAFKR NHLKEVKGSSTLSQITF
NAFDQNDGDKRFGKKVVVRTHNNSHMLADGERFIIDPKLSSTMVDLEKVL

5 KIEGLDYSTLRQTTQTQFREMTTAGKALLSKSNLRQGEKQKFLQEAQFFLGR
VDLDKAIKA EKALVTKKATKNGHLLERSINKAVLAYNNSAIKKANVKRLEK
ELDLLTDLVEGKGPLAQATMVQGVYLLKTPLPLPEYYIGLNVYFDKSGKLIYA
LDMSDTIGEGQKDAYGNPILNVDEDNEGYHTLAVATLADYEGLYIKDILNSSL
DKIKAIRQIPLAKYHRLGIFQAIRNAAAEADRLLPKTPKGYLNEVPNYRKKQM
EKNLKPVDYKTPIFNKALPNEKVDGDRAAKGHNINAETNNSVAVTPIRSEQQL
HKSQSDVNLPQTSSKNNFIYEILGYVSLCLLFLVTAGKKKGKRARK*

10 Sequence description:

A] Length 3168 bp - 1056 aa (Partial sequence)
B] Obvious signal peptide with Shine Dalgarno
sequence upstream of the ATG start codon.

15

ID-88

20 Clone RS-56

(SEQ ID NO: 25)
GCAGGATACATCATGCACAAGCACGAGGCTATCGTGTCATGCTGGGGGTCA
ACCCAGGAAGACATGTCGGCACAAGCTGAAGATTTCTTTACAGTCTGTACA
25 CAATAAAGAGACGGGTAAGAGCGCTTTTAATGACAAAGAACGACTAGCAA
TT

(SEQ ID NO: 26)
AGYIMHKHEAIVSCWGQPRKTCRHKLKISLQSVHNKETGKSAFNDKERLAI
30

Sequence description:

35 A] Length:153 bp - 51 aa (partial sequence)
B] No signal peptide visible, insufficient
sequence data to determine the presence of a
Shine Dalgarno sequence.

40 ID-89

Clone RS-58

(SEQ ID NO: 27)

GTGTCATTTATGCAAAGAAAATCCTATTTAAAATCCATGAGTGTTCTTACT
TTAACAGCTTGTCTTATATCAGGATATGTGGTTAAAGATATTGCTATGTTA
CATGCAGTATCTGCCAGTGAGAAGAAAGCAAATAATGTCAGTCCGAGAGA
5 AAATCTCTACAGGGCTGTCAATGATAATTGGCTAGCCAATACAAAACCTCA
AACAAGGGCAGACTAGTGTTAATAGTTTTTCAGAAATTGAGGATAAATTA
AAGCAACTGTTAGTGTCTGATATGGCTAAAATGGCCTCAGGAAAGATTGA
AACAACCAATGATGAACAGAAAAAAATGGTTGCATACTATAAACAAGGTA
TGGACTTTAAAACAAGAGATAAAAAATGGTCTCAAACCTCTAAAACCAAGTT
10 TTACAAAAAAGTTGAAGCAGTCTCTTCAATGAAAGACTTTCAAAGTTTGGCC
CATGATTTTGTGATGAGTGGTTTTGTTTTACCATTTGGTTTGACTGTGGAAA
CCAATGCTCGAGATAATAGCCAAAAGCAATTGGTGCTTCGTCAAGCACCC
GCATTACTTGAATCACCTGACCAATATAAGAAGGGCAATAAAGAAGGTGA
GGCTAAATTATCAGCTTACCGTACTTCAGCAATGGCTTTGCTTAAACAAGC
15 TGGAAAAAGTAACATTGAAGATAGAAAACCTAGTTAAACAAGCTATAGCAT
TTGATAGACTCTTATCAGAAAAAACGCAAGTTGATCAAAGTAAAATCACA
GCTGAAAGTGAGACAGCTGCGGGGCGATATAACCCTGAAAGTATGGAAAC
GGTTCACAATTACGCCAAGGAATTTGACTTTAAAGAATTGATTGAAAAACT
AGTTGGGCCAACGAATAAGGCAGTCAATGTAGAAGATAAACTTATTTTA
20 AACAGGTTAATGATGTTATAAATAGTAAACAATTAGCCAATATGAAAGCA
TGGATGATGATTTCTATGCTAGTTGATCAATCAGATTTTCTAGGAGAACAA
AATCGTCAAGCAGCGAGTGCTTTTAAGAATGTTGCGTCTGGTTTGACTCAG
ATTGAATCGAAAGAAAAAATGCTTACACCCAATTAG

25

(SEQ ID NO: 28)

MSFMQRKSYLKSMVLTLTACLISGYVVKDIAMHLHAVSASEKKANNVSPREN
LYRAVNDNLWLANTKLKQGQTSVNSFSEIEDKLKQLLVSDMAKMASGKIETTN
DEQKKMVAYYKQGMDFKTRDKNGLKPLKPVLPQKLEAVSSMKDFQSLAHDF
30 VMSGFVLPFGLTVETNARDNSQKQLVLRQAPALLES PDQYKKGNKEGEAKLS
AYRTSAMALLKQAGKSNIEDRKLVKQAIAFDRLLESEKTQVDQSKITAESETAA
GRYNPESMETVHNYAKEFDKELIEKLVGPTNKA VNVEDKTYFKQVNDVINS
KQLANMKAWMMISMLVDQSDFLGEQNRQAASAFKNVASGLTQIESKEKMLT
PN*

35

Sequence description:

- 40 A] Length: 1095 bp - 365 aa (full-length gene)
B] an GTG (possible ATG start codon located 7 bp
further downstream) start codon with an obvious
signal peptide. Shine Dalgarno sequence present
upstream of the ORF.

ID-90

Clone RS-59

5

(SEQ ID NO: 29)

ATGGAAATGCCTAAAAGAAATGAATTACTCAATAAAGAAATTAAAATGAG
TATTGATAAACTTAGATATAAAGAACCAGAGAGTGAACATGACAAGCGAC
CTACTTTTTATTTGGTAGTACTTATACTTGTTACTGTAGCAGTTATATTGTC
GTTATTTAAATATTTTTTATAG

10

(SEQ ID NO: 30)

MEMPKRNELLNKEIKMSIDKLRYKEPESEHDKRPTFYLVVLILVTVAVILSLFK
YFL*

15

Sequence description:

20

A] Length: 174 bp - 58 aa(full length gene)
B] No obvious signal peptide, but Shine
Dalgarno sequence is present upstream of ATG
start codon.

25

ID-91

Clone RS-62 (partial sequence)

30

(SEQ ID NO: 31)

ATGCAGGTATTTTTTAAATATTGTCAATAAATTCTTTGATCCAGTTATTCATA
TGGGTTCGGGAGTTGTGATGCTAATTGTCATGACAGGTTTAGCCATGATAT
TTGGAGTGAAGTTTTCTAAAGCACTTGAAGGTGGTAT

35

(SEQ ID NO: 32)

MQVFLNIVNKFFDPVIHMGSGVVMLIVMTGLAMIFGVKFSKALEGG

Sequence description:

40

A] Length: 141 bp - 41 aa (partial sequence
B] Shine Dalgarno sequence present upstream of
ATG start codon with a possible signal peptide
present



ID-92

5 Clone RS-69 (partial sequence)

(SEQ ID NO: 33)

10 ATGAAAAAGAAAACATTCAGTGCTTATAACTTTTTTAACGGCTCTTATCCTT
TGTCTTTTGACAGTGCTTTTTTATCTTTCCATTTTATTGGATTATGACAGGAG
CTTTTAA

(SEQ ID NO: 34)

MKKKTFSAYNFLTALILCLLTVLFIFPFYWIMTGAF

15

Sequence description:

20 A] Length: 110 bp -36 aa (Partial sequence)
B] Possible signal peptide with Shine Dalgarno
sequence directly upstream of the ATG start
codon.

25 ID-93

Clone RS-70

(SEQ ID NO: 35)

30 ATGACTGAGAACTGGTTACATACTAAAGATGGTTCAGATATTTATTATCGT
GTCGTTGGTCAAGGTCAACCGATTGTTTTTTTACATGGCAATAGCTTAAGT
AGTCGCTATTTTGATAAGCAAATAGCATATTTTTCTAAGTATTACCAAGTT
ATTGTTATGGATAGTAGAGGGCATGGCAAAGTCATGCAAAGCTAAATAC
35 CATTAGTTTCAGGCAAATAGCAGTTGACTTAAAGGATATCTTAGTTTCATTT
AGAGATTGATAAAGTTATATTGGTAGGCCATAGCGATGGTGCTAATTTAGC
TTTAGTTTTTCAAACGATGTTTCCAGATATGGTTAGAGGGGCTTTTGCTTAAT
TCAGGGAACCTGACTATTCATGGTCAGCGATGGTGGGATATTCTTTTAGTA
AGGATTGCCTATAAATTCCTTCACTATTTAGGGGAACTCTTCCGTATATG
AGGCAAAAAGCTCAAGTTATTTGCTTATGTTGGAGGATTGAAGATTAGT
40 CCAGCTGATTTACAGCATGTGTCAACTCCTGTAATGGTTTTGGTTGGAAAT
AAGGACATAATTAAGTTAAATCATTCTAAGAACTTGCTTCTTATTTTCCA
AGGGGGGAGTTTTATTCTTTAGTTGGCTTTGGGCATCACATTATTAAGCAA
GATTCCCATGTTTTTAATATTATTGCAAAAAAGTTTATCAACGATACGTTG
AAAGGAGAAATTGTTGAAAAAGCTAATTGA

(SEQ ID NO: 36)

MTENWLHTKDGSDIYYRVVGQGQPIVFLHGNSLSSRYFDKQLAYFSKYYQVIV
MDSRGHGKSHAKLNTISFRQIAVDLKDILVHLEIDKVILVGHSDGANLALVFQ
5 TMFPDMVRGLLLNSGNLTIHGQRWWDILLVRIAYKFLHYLGKLFYPMRQKA
QVISLMLEDLKISPADLQHVSTPVMVLVGNKDIKLNHKKLASYPFRGEFYSL
VGFGHHIHKQDSHVFNIIAKKFINDTLKGEIVEKAN*

10 Sequence description:

A] Length: 744 bp - 248 aa (full length gene)
B] No obvious signal peptide, but Shine
Dalgarno sequence upstream of the ATG start
15 codon.

ID-94

20 Clone RS-71

(SEQ ID NO: 37)

ATGGTAGCAAAAAGAGTTAGGTAAAAATAGCTTTACTATCCCAACTATTTGT
TCTAATTGCTCCGCAGGTACTGCCATTGCAGTTGTATATAATGATGACCAT
25 TCTTTCTTAAGATACGGCTATCCCGAGTCTCCACTTCATATTTTTATCAATA
CACGGATCATTGCACAGGCACCAAGCAAATATTTTTGGGCTGGTATTGGGG
ACGGTATTTCAAAAAGCCCCTGAAGTAGAACGTGCTACCTTAGAGGCTAAG
ACCAATAAACTACCACATACTGCAGTGTTAGGACAAGCAGTCGCTCTGTCT
TCAAAGGAAGCTTTTTATCAATTTGGTGAACAAGGTCTAAAAGACGTTGAA
30 GCTAATTTAGCTTCGCGTGCAGTTGAAGAAATTGCGCTTGATATCTTA

(SEQ ID NO: 38)

MVAKELGKNSFTIPTICSNCSAGTAIAVVYNDDHSFLRYGYPEsplHIFINTRIIA
QAPSKYFWAGIGDGISKapeVERATLEAKTNKLPHTAVLGQAVALSskeAFY
35 QFGEQGLKDVEANLASRAVEEIALDIL

Sequence description:

40

A] Length: 405 bp - 135 aa (Partial sequence)
B] No obvious Shine Dalgarno sequence upstream
of the ATG start codon, probable signal
peptide present at the N-terminus.

ID-95

Clone RS-73

5 (SEQ ID NO: 39)
TTGAGGGAACTTACTGGAAAATTTCAAGCGATTGCGATAAAATAAATCTT
GCAGAGTTTTCTAGAGAAAGGAGGTCAGATTTATTGGAGTGGCAAGATCT
AGCGCAGTTACCTGTATCTATTTTTTAAAGACTATGTTACAGATGCTCAAGA
CGCGGAAAAACCTTTTATATGGACAGAAGTATTTTTTAAGGGAGATTAATCG
10 CTCAAATCAAGAAATTATTTTGCATATTTGGCCGATGACTAAGACAGTCAT
TCTGGGGATGTTAGATCGAGAATTACCACATTTAGAATTAGCTAAAAAAG
AAATCATCAGTCGTGGTTATGAACCAGTTGTTTCGGAATTTTGGAGGTCTCG
CAGTTGTAGCTGATGAAGGAATTTTAAATTTTTCATTGGTTATTCCAGATGT
TTTTGAGAGAAAATTGTCTATCTCAGATGGGTATCTTATAATGGTCGATTTT
15 ATTAGAAGTATATTTTCGGATTTTATCAACCTATTGAGCACTTTGAAGTA
GAGACCTCCTATTGTCCTGGTAAGTTTGATCTTAGTATAAATGGCAAAAAA
TTTGCTGGCTTGGCTCAGCGCCGTATAAAGAATGGTATTGCGGTATCAATT
TACCTTAGCGTTTGTGGCGATCAAAAAGGGCGGAGTCAAATGATTTTCAGAT
TTTTATAAGATTGGTCTAGGTGATACGGGTAGTCCAATTGCTTATCCAAAT
20 GTAGATCCTGAAATTATGGCTAATCTATCTGATCTATTAGATTGTCCTATG
ACAGTAGAAGATGTTATTGATCGTATGTTGATTAGCCTTAAACAAGTAGGT
TTTAATGATCGTTTACTGATGATTAGACCCGATTTAGTTGCAGAGTTTGAT
AGATTTTCAGGCTAAGTCTATGGCTAATAAGGGGATGGTGAGCAGAGATGA
ATAA

25 (SEQ ID NO: 40)
MRETYWKISSDCDKINLAEFSRERRSDLLEWQDLAQLPVSIFKDYVTDAQDAE
KPFIWTEVFLREINRSNQEILHIWPMTKTVILGMLDRELPHLELAKKEIISRGYE
PVVRNFGGLAVVADEGILNFSLVIPDVFERKLSISDGYLMVDFIRSIFSDFYQPI
30 EHFEVETSYCPGKFDLSINGKKFAGLAQRRIKNGIAVSIYLSVCGDQKGRSQMI
SDFYKIGLGD TGSP IAYPNVDPEIMANLSDLLDCPMTVEDVIDRMLISLKQVGF
NDRLLMIRPDLVAEFDRFQAKSMANKGMVSRDE*

35 Sequence description:

40 A] Length: 921 bp -307 aa (Full-length gene sequence)
B] No obvious Shine Dalgarno sequence upstream
of the TTG start codon or signal peptide
visible. Actual start point may be a further
85 bp downstream (TTG). This start point is
preceded by a typical Shine-Dalgarno sequence.

ID-96

Clone RS-74

5

(SEQ ID NO: 41)

TTGGAAGGTTTACTTATTGCATTGATTCCCATGTTTGCGTGGGAAAGTATT
GGATTTGTTAGTAATAAAAATTGGAGGGGCGTCCAAATCAACAAACATTTGG
AATGACTTTAGGAGCATTGCTATTTGCGATTATCGTATGGTTATTTAAACA
10 GCCAGAGATGACTGCCTCATTGTGGATTTTTGGTATCTTAGGTGGTATCCT
ATGGTCAGTCGGCCAAAATGGTCAATTTCAAGCAATGAAATATATGGGAG
TCTCTGTTGCTAATCCACTGTCAAGTGGTGCACAATTAGTAGGTGGAAGCC
TAGTTGGTGCTTTAGTCTTTCATGAATGGACTAAGCCAATCCAATTTATTTT
AGGATTGACAGCGTTGACATTATTAGTTATCGGCTTCTATTTCTCAAGTAA
15 ACGTGATGTTTCAGAACAAGCTTTGGCAACACATCAAGAGTTTTCAAAG
GATTTGCTACAATTGCTTATTCAACTGTAGGTACATCTCGTACGCAGTTTT
ATTTAACAACATTATGAAGTTCGACGCTATGGCCGTCATTTTACCCATGGC
TGTTGGAATGTGTCTAGGTGCAATTTGTTTCATGAAGTTTCGTGTAACTTT
GAGGCTGTTGTTGTTAAAAATATGATTACAGGTCTCATGTGGGGCGTTGGT
20 AATGTCTTCATGTTATTGGCAGCAGCTAAAGCAGGGCTAGCAATTGCTTTT
AGTTTTTCTCAACTTGGAGTAATTATCTCTATTATTGGTGGTATTTTATTTT
AGGTGAGACAAAAACGAAGAAAGAGCAGAAATGGGTGTCATGGGTATC
CTTTGTTTTGTTATGGGTGCTATATTACTTGGTATTGTTAAATCTTATTAA

25

(SEQ ID NO: 42)

MEGLLIALIPMFAWESIGFVSNKIGGRPNQQTFGMTLGALLFAIIVWLFKQPEM
TASLWIFGILGGILWSVGQNGQFQAMKYMGSVSVANPLSSGAQLVGGSLVGAL
VFHEWTKPIQFILGLTALTLLVIGFYFSSKRDVSEQALATHQEFSKGFATIAYST
VGYISYAVLFNNIMKFDAMAVILPMAVGMCLGAICFMKFRVNFEAVVVKNMI
30 TGLMWGVGNVFMLLAAAKAGLAIAFSFSQLGVIISIIGGILFLGETKTKKEQK
WVVMGILCFVMGAILLGIVKSY*

35

Sequence description:

A] Length: 867 bp - 289 aa (full-length gene)
B] Possible Shine Dalgarno sequence upstream of
GTG start codon, no obvious signal peptide
present.

40

ID-97

Clone RS-75

(SEQ ID NO: 43)

5 ATGACAACCTTACTACGAAGCTATAAACTGGAACGAAATTGAAGATGTTAT
TGATAAATCAACTTGGGAAAACTAACCGAACAATTTTGGCTCGATACAC
GTATCCCTTTATCAAATGACTTAGACGATTGGCGCAAACCTTTCCGCTCAAG
AAAAAGATCTTGTTGGCAAGGTTTTTGGAGGCTTAACCCTACTTGATACCA
TGCAATCAGAACTGGTGTGGAAGCTATTCGTGCCGATGTTTCGCACGCCTC
ACGAAGAAGCTGTCTTAAACAATATTCAATTCATGGAATCTGTTACGCTA
AATCTTATTCTTCAATTTTCTCAACTTTAAATACTAAATCAGAAATTGAAG
10 AAATTTTCGAGTGGACTAATAATAATGAGTTCCTTCAAGAAAAAAGCACGT
ATTATCAATGACATTTATGCTAATGGAAATGCCCTTCAAAAAAAGGTGGCT
TCCACCTACCTCGAACTTTCCTTTTTTTATTCTGGCTTTTTTCACACCTCTTTA
CTATTTGGGAAATAATAAGTTAGCAAATGTTGCTGAAATCATTAAATTAAT
TATTCGTGATGAATCTGTACATGGTACTTATATCGGTTACAAATTCCAGCTT
15 GGTTTTAACGAATTACCAGAAGATGAGCAAGAGAATTTTCGTGATTGGAT
GTATGACCTCCTTTATCAGCTGTATGAAAACGAAGAAAAATACACCAAGA
CACTTTATGATGGCGTAGGATGGACTGAAGAAGTTATGACCTTTTTACGCT
ACAATGCTAATAAAGCTCTTATGAATTTAGGACAAGATCCTTTATTCCCAG
ATACAGCAAATGATGTCAACCCAATTGTTATGAATGGTATTTCAACAGGAA
20 CATCAAACCATGACTTCTTCTCTCAAGTAGGTAATGGTTACCTACTTGGTA
GCGTTGAAGCTATGCATGATGATGACTATAACTATGGATTATAA

(SEQ ID NO: 44)

25 MTTYEAINWNEIEDVIDKSTWEKLTEQFWLDTRIPLSNDLDDWRKLSAQEK
DLVGKVFGGLTLLDTMQSETGVEAIRADV RTPHEEA VLNNIQFMESVHAKSY
SSIFSTLNTKSEIEEIFEWTNNNEFLQEKARIINDIYANGNALQKKVASTYLETF
LFYSGFFTPLYLGNKLANVAEIILIRDES VHGTYIGYKFQLGFNELPEDEQ
ENFRDWMYDLLYQLYENEEKYTKTLYDGVGWTEEVMTFLRYNANKALMNL
30 GQDPLFPDTANDVNPIVMNGISTGTSNHDFFSQVGNGYLLGSVEAMHDDDDYN
YGL*

Sequence description:

35 A] Length: 960 bp - 320 aa (full length gene)
B] Shine Dalgarno sequence present upstream of
ATG start codon, but no signal peptide
present.

40

ID-98

Clone RS-77 (partial sequence)

(SEQ ID NO: 45)

ATGAATTGGTCACGTATCTGGGAACTCGTAAAAATTAATATCCTTTATTCA
AACCCCTCAGACTCTATCGGCACTAAGAAAAAAGCAAGAAAAGCATCCTAA
AAAAGAATTTTCAGCTTATAAATCCATGTTTAGAAATCAGTTATTTTCAGAT
5 TTTGCTCTTTTCAATAATTTATGTATTTCTCTTTGTATCACTTGATTTTAAAG
AATATCCGGGCTATTTACGTTCTACATTGGTATCTTTACACTAGTATCCAT
TATCTACTCTTTTATTGCGATGTACAGTGTTTTCTATGAGAGTGACGATGTT
AA

(SEQ ID NO: 46)

MNWSRIWELVKINILYSNPQTLALRKKQEKHPKKEFSAYKSMFRNQLFQILL
FSIIYVFLFVSLDFKEYPGYFTFYIGIFTLVSIIYSFIAMYSVFYESDDV

Sequence description:

- A] Length: 311 bp - 103 aa (Partial sequence)
- B] Shine Dalgarno sequence present upstream of
ATG start codon, no obvious signal peptide at
N-terminus.

ID-99

Clone RS-78 (partial sequence)

(SEQ ID NO: 47)

TAATCTTTTAGTCAACGGAGCAACAGGAAAATTGCAGGCTATGCGACAGA
TATTCACCACATAATTTAGCAGAAGTCATTGATGCTGTCGTGTACATGAT
30 TGATCACCCCTAAAGCTAAATTAGATAAATTAATGGAATTTCTACCTGGTCC
AGATTTTCCAACCTGGCGCTATCATTCAAGGAAAAGATGAAATTCGTAAGG
CATATGAGACTGGTAAGGGGAGAGTAGCGGTTCGCTCGCGAACTGCTATT
GAAACCTTAAAAGGTGGTAAGAAACAAATTATTGTTACTGAAATTCCTTAT
GAAGTTAAT

(SEQ ID NO: 48)

SFSQRSNRKIAGYATDIPPHNLAVIDAVVYMHDPKAKLDKLMEFLPGPDFPT
GAIQGKDEIRKAYETGKGRVAVRSRTAETLKGGKKQIIVTEIPYEVN

Sequence description:

- A] Length: 312 bp - 104 aa (Partial sequence)
- B] No obvious Shine Dalgarno sequence or a

signal peptide. Both N- and C- termini of ORF
yet to be elucidated.

5 ID-100

Clone RS-79

(SEQ ID NO: 49)

10 ATGGGACGTAAGTGGGCCAATATTGTTGCCAAAAAGACTGCTAAAGATGG
TGCTAACTCAAAAAGTATACGCTAAATTCGGTGTTGAAATATATGTTGCTGC
AAAGCAAGGTGAACCAGACCCCGAGTCAAACCTCAGCTCTAAAATTCGTTT
TGGACCGTGCTAAGCAAGCACAAAGTTCCAAAGCATGTTATTGATAAAGCG
ATTGATAAAGCCAAAGGAAACACAGATGAAACTTTCGTAGAGGGACGCTA
15 TGAAGGTTTTTGGTCCAAATGGTTCAATGATTATTGTGGATACTTTGACATC
AAATGTTAACCGTACGGCAGCAAATGTACGTACTGCTTACGGTAAGAACG
GTGGCAATATGGGAGCTTCAGGATCGGTATCCTACTTATTTGATAAAAAAG
GTGTCATCGTTTTTTGCTGGTGATGATGCTGACACTGTCTTCGAACAATTACT
TGAAGCGGATGTAGACGTAGATGATGTTGAAGCAGAAGAGGGAACAATA
20 ACAGTTTATACCGCCCCAACAGATCTTCATAAAGGTATCCAAGCACTTCGC
GATAATGGTGTAGAAGAATTCCAAGTTACTGAACTTGAAATGATTCCTCAA
TCAGAAGTAGTATTGGAAGGTGATGACCTTGAAACTTTTGAAAAGCTT

(SEQ ID NO: 50)

25 MGRKWANIVAKKTAKDGAN SKVYAKFGVEIYVAAKQGEPDPESNSALKFVL
DRAKQAQVPKHVIDKAIDKAKGNTDETFVEGRYEGFGPNGSMIIVDTLTSNV
NRTAANVRTAYGKNNGNMGASGSVSYLFDKKGVIVFAGDDADTVFEQLLEA
DVDVDDVEAEEGTITVYTAPDLHKG IQALRDNGVEEFQVTELEMIPQSEVVL
EGDDLETFEKL

30

Sequence description:

35 A] Length: 654 bp - 218 aa (Partial sequence)
B] Possible Shine Dalgarno sequence upstream
of ATG start, no obvious signal peptide

40 ID-101

Clone RS-80

(SEQ ID NO: 51)

TTGGAGAAATATTTGAAGAACCCGATTACATGGATTGGATTAGTTCTTGTG
GTTACGTGGTTTTTAACTAAAAGTAGTGAATTTTGGATTGTTGGTGTGTGTG
TCTTGTTGTTAGTATTTGCTAGTCAAAGTGAT

5

(SEQ ID NO: 52)

MEKYLKNPITWIGLVLVVTWFLTKSSEFLIFGVCVLLLVFASQSD

10 Sequence description:

A] Length: 135 bp - 45 aa (partial sequence)

B] Shine Dalgarno sequence upstream of TTG
start codon with possible signal peptide

15

evident at N-terminus.

ID-102

20

Clone RS-81

(SEQ ID NO: 53)

ATGACACAATCAGATGCATATCTCTCGTTGAACGCGAAGACACGCTTTAGA
GATCGCACAGGTAATTATCATTCTTACTTCGGATAAAGAGGCTGTTGAACAA
TATATGATAGAACATGTTGAACCTAATACGATGGTGTTTACATCACTAATT
GAAAAGCTAGATTATTTGGTTTCTAATAACTACTATGAATCGGACCTTCTA
AAACAATATAACCTTGAGTTTATTTGCCAAATTTTGGAGCATGCATACGCT
AAGAAATTTGCTTTTCTAAATTTTATGGGGGCTTTAAAATTTTATAATGCTT
ATGCTCTTAAT

30

(SEQ ID NO: 54)

MTQSDAYLSLNAKTRFRDRTGNYHFTSDKEAVEQYMIEHVEPNTMVFTSLIE
KLDYLVSNYYESDLLKQYNLEFICQIFEHAYAKKFAFLNFMGALKFYNAYA
LN

35

Sequence description:

40

A] Length: 318 bp - 106 aa (Partial sequence)

B] Shine Dalgarno sequence present upstream of
ATG start codon, no obvious signal peptide

ID-103

Clone 2-11A

5 (SEQ ID NO: 55)
ATGGTATTTATGGCAAATAAGAAAAAAAAACAAAAGGAAAGAAAACCAGAA
GACCTACTAAGGCAGAAATAGAGCGTCAAAGAGCTATTCAAAGGATGATT
ACTGCTCTTGTTTTAACAATTATTCTCTTCTTTGGTATTATCAGATTAGGTA
TTTTTGGTATTACAGTCTATAACGTCATCCGTTTTATGGTAGGTAGCTTGGC
10 TTACTTATTTATTGCGGCAACTTTAATCTACCTTTATTTCTTTAAATGGTTG
CGAAAGAAAGATAGCTTAGTAGCAGGTTTTTTGATAGCTTCTTTAGGATTA
TTGATTGAGTGGCATGCTTACCTTTTCTCAATGCCTATTTTGAAAGATAAA
GAAATTTTGCGTTCAACTGCTCGATTAATTGTGTCTGATTTAATGCAATTTA
AAATCACTGTTTTTTGCCGGTGGAGGTATGTTGGGTGCTTTGATTTACAAGC
15 CAATTGCTTTTCTCTTTTCTAATATTGGTGCCTATATGATTGGTGTCTCTTC
ATCATTTTGGGTCTCTTTTTAATGAGTTCTCTGGAAGTTTATGACATCGTCG
AATTTATTAGAGCTTTTAAAAATAAAGTGGCAGAGAAGCACGAGCAAAAT
AAAAAGGAGCGTTTTTGCTAAGCGAGAGATGAAAAAAGCAATCGCTGAACA
AGAGCGCATAGAGCGTCAAAAAGCTGAAGAAGAAGCTTATTTAGCTTCGG
20 TTAATGTAGACCCTGAAACGGGTGAGATTCTAGAGGATCAAGCTGAGGAC
AATTTGGATGATGCGCTACCACCTGAGGTAAGTGAAACATCAACTCCGGT
ATTTGAGCCAGAGATCCTTGCTTATGAGACATCGCCTCAAAATGATCCTTT
ACCAGTAGAGCCGACAATTTATTTAGAAGACTATGATTTCGCCGATTCCTAA
TATGAGAGAAAATGATGAGGAAATGGTTTATGATTTAGATGATGATGTAG
25 ATGATAGTGATATAGAAAATGTCGACTTTACACCTAAAACGACACTGGTTT
ATAAATTACCAACGATAGATTTATTTGCACCAGATAAGCCTAAAAATCAAT
CCAAAGAAAAGGATTTAGTCCGAAAGAATATCAGAGTTTTTAGAAGAAACA
TTTAGAAGTTTTTGGTATCGATGTAAAAGTAGAACGTGCTGAAATTGGACCA
TCAGTTACTAAATATGAAATTAAACCAGCAGTTGGAGTTCGTGTGAATCGT
30 ATTTCAAATCTATCTGACGACCTAGCTCTTGCTCTTGCAGCAAAAAGATGTG
CGTATAGAAGCACCAATTCCTGGAAAATCATTAAATAGGTATTGAAGTTCCT
AACTCAGAAATTGCAACGGTTTCTTTCCGCGAACTTTGGGAACAATCTGAT
GCCAATCCTGAAAACCTTTTAGAAGTACCCTAGGAAAAGCTGTTAACGG
CAATGCTCGCAGTTTTAACTTAGCTAGAATGCCGCATCTTTTGGTAGCTGG
35 TTCAACTGGTTCAGGTAAATCTGTGGCAGTTAATGGAATTATTTCAAGTAT
TTTGATGAAGGCACGTCCAGATCAAGTTAAGTTTATGATGATTGATCCCAA
AATGGTTGAATTATCTGTTTATAATGATATTCCACATTTATTAATCCCTGTT
GTAACCAATCCGCGTAAAGCAAGTAAGGCACTCCAAAAAGTTGTTGATGA
AATGGAAAATCGATACGAGTTATTTAGCAAAATTGGTGTGCGTAATATAG
40 CAGGTTATAATACAAAGGTTGAAGAGTTTAATGCTTCCTCTGAGCAAAAAC
AAATGCCTTTGCCTTTAATCGTTGTCATTGTAGATGAATTGGCTGACTTGAT
GATGGTTGCTAGTAAAGAAGTTGAAGATGCTATTATTCGTTTGGGGCAAAA
AGCACGTGCTGCAGGTATCCATATGATTCTTGCAACTCAACGTCCATCCGT
AGATGTTATTTCTGGTTTGATTAAAGCAAATGTTCCGTCGCGTATTGCATTT

5 GCTGTTTCAAGTGGTACTGATAGCCGTACGATCCTTGATGAAAATGGTGCT
GAAAAGCTCTTGGGACGGGGTGACATGCTCTTTAAGCCTATTGATGAGAAT
CATCCAGTACGACTACAAGGTTCTTTATTTTCAGATGATGATGTTGAAAGG
ATCGTTGGTTTTATCAAAGACCAAGCCGAGGCTGACTATGATGATGCCTTT
10 GATCCTGGAGAAGTATCTGAAACAGATAACGGCTCTGGTGGTGGCGGCGG
AGTACCTGAAAGTGATCCTCTTTTTGAAGAAGCCAAGGGACTCGTTTTAGA
GACGCAAAAAGCAAGTGCCTCAATGATTCAACGCCGATTGTCTGTTGGTTT
CAATAGAGCAACAAGACTAATGGAAGAATTAGAAGCAGCGGGGGTTATTG
GTCCAGCAGAAGGAACCAAGCCACGAAAAGTTTTAATGACTCCAACCTCCG
AGTGAATAA

(SEQ ID NO: 56)
15 MVFMANKKKTKGKKTRRPTKAEIERQRAIQRMITALVLTILFFGIIRLGIFGIT
VYNVIRFMVGSLAYLFIAATLIYLYFFKWLRKKDSL VAGFLIASLGLLIEWHA
YLFSMPILKDKEILRSTARLIVSDLMQFKITVFAGGGMLGALIYKPIAFLFSNIG
AYMIGVLFILGLFLMSSLEVYDIVEFIRAFKNKVAEKHEQNKKERFAKREMK
KAIAEQERIERQKAEEEEAYLASVNVDPETGEILEDQAEDNLDDALPPEVSETST
PVFEPEILAYETSPQNDPLPVEPTIYLEDYDSPIPNMRENDEEMVYDLDDD VDD
20 SDIENVDFTPKTTLVYKLPTIDLFAPDKPKNQSKEKDLVRKNIRVLEETFRSFGI
DVKVERAEIGPSVTKEYEIKPAVGVRVNRISNLSDDLALALAAKDVRIEAPIPGK
SLIGIEVPNSEIATVSFRELWEQSDANPENLLEVPLGKAVNGNARSFNLARMPH
LLVAGSTGSGKSVAVNGISSLMKARPDQVKFMMIDPKMVELSVYNDIPHLLI
PVVTNPRKASKALQKVVDENRYELFSKIGVRNIAGYNTKVEEFNASSEQK
25 QMPLPLIVVIVDELADLMMVASKEVEDAIIRLGQKARAAGIHMILATQRPSVD
VISGLIKANVPSRIAFAVSSGTD SRTILDENGAEKLLGRGDMLFKPIDENHPVRL
QGSFISDDDVERIVGFIKDQAEADYDDAFDPGEVSETDNGSGGGGGVPESDPL
FEEAKGLVLETQKASASMIQRRLSVGFNRATRLMEELEAAGVIGPAEGTKPRK
VLMTPTPSE*

30 Sequence description:

- 35 A] Length: 2451 bp - 817 aa (Full-length gene)
B] Shine Dalgarno sequence present upstream of
ATG start codon, possesses a potential signal
peptide

40 ID-104

Clone 2-18/22b

(SEQ ID NO: 57)

ATGTCACAAGAGCAAGGAAAAATTTATATTGTAGAAGATGATATGACGAT
TGTGTCACTTTTTAAAAGATCATTTATCAGCTAGCTATCATGTCTCTAGTGTC
AGCAATTTTCGTGATGTGAAACAAGAAATTATCGCATTTC AACCCCGATTG
ATACTAATGGATATTACGTTACCCTATTTTAATGGTTTTTACTGGACTGCAG
5 AATTGCGTAAGTTTTTAACAATTCCTATTATTTTCATTTCATCTAGTAATGA
TGAAATGGATATGGTTATGGCATTAAATATGGGGGGGTGATGACTTTATTTC
AAAACCATTCTCTCTAGCTGTATTAGATGCTAAGCTAACTGCTATTTTAAG
GAGAAGTCAACAATTTATCCAACAGGAATTAACTTTTGGGGGGATTACGTT
GACAAGAGAAGGGTTATTGTCTAGCCAAGATAAAGAGGTTATTTTATCGC
10 CAACAGAAAATAAAAATCCTATCTATCTTGCTCATGCATCCTAAACAAGTAG
TCTCAAAAGAGTCTCTATTAGAGAAACTTTGGGAAAATGATAGTTTTATTG
ATCAAAATACACTTAATGTTAATATGACACGCTTACGTAAAAAAATTGTCC
CAATAGGTTTTGATTACATTCATACAGTGAGAGGAGTTGGGTATTTACTAC
AATGA

15 (SEQ ID NO: 58)
MSQEQGKIYIVEDDMTIVSLLKDHLASASYHVSSVSNFRDVKQEIIAFQPDILM
DITLPYFNGFYWTAELRKFLTPIIFISSNDEMMDMVMALNMGGDDFISKPFSLA
VLD AKLTAILRRSQQFIQQELTFGGFTLTREGLLSSQDKEVILSPTENKILSILLM
20 HPKQVVSKESSLLEKLWENDSFIDQNTLNVNMTRLRKKIVPIGFDYIHTVRGVG
YLLQ*

Sequence description:
25 A] Length: 669 bp - 223 aa (full-length gene
sequence)
B] Shine Dalgarno sequence present upstream of a GTG start codon.
Was not identified directly by LEEP. This gene was found upstream of
30 gene ID-10 described in WO 00/06736.

ID-105

35 Clone 2-20

(SEQ ID NO: 59)
ATGTATCAAACCTCAGACAAATAAGGAAAAAATTTGTTTTATTTTTGAAATTA
TTTATCCCAGTATTFGATTTATCAATTTGCTAATTTTTCAGCTACTTTTATTGA
40 TTCGGTTATGACTGGACAGTATAGTCAGCTACATTTGGCAGGTGTGTCAAC
TGCTAGTAATTTATGGACTCCGTTTTTCGCTTTATTAGTAGGTATGATTTCA
GCATTAGTACCAGTAGTTGGTCAACATTTGGGTAGAGGAAATAAAGAACA
AATTCGCACAGAATTTTCATCAATTTCTATATTTAGGTTTGATACTGTCCTTA
ATATTATTTTTAATCATGCAATTTATTGCTCAACCTGTCTTGGGGAGTTTGG

5 GTTTAGAAGATGAAGTTCTAGCAGTTGGTCGTGGTTATTTAAATTATATGT
TGATTGGAATCATGCCGCTGGTGTGTTTAGCATTTGCCGTTCAATTCTTTGA
TGCATTGGGGTTAACAAGGTTATCTATGTATCTGATGCTTTTAATTCTACCC
TTTAATTCATTTTTTTAATTATATGCTTATCTACGGTAAATTTGGTATGCCTA
GACTAGGAGGTGCGGGGGCAGGTCTTGGAACCTTCTTTAACTTATTGGGCTA
TTTTTATTGGTATTATTATTGTGATGTCACCTCATCCTCAAATTA AACATA
TCATATATGGACTCTGGAAAGAATAAAAGCTCCTTTGATTATTGAAGATAT
TCGATTGGGATTACCGATTGGTTTACAAATTTTTGCAGAAGTTGCAATTTTT
GCAGTAGTAGGCTTATTCATGGCAAAATTTTCTTCAATCATTATTGCAGCA
10 CATCAGGCTGCTATGAATTTTTTCATCATTAAATGTATGCATTTCTTTAAGTA
TTTCCACTGCTCTAGCTATTACAATATCGTTTGAAGTAGGGGGCAGAGCGCT
TTCAGGACGCAACCACTTATAGTAGGATAGGACGCTTAACAGCGGTAGGG
ATTACATCAGGAACCTTACTATTTTTATTTCTATTTCTGTGAGAATGTAGCAG
CAATGTATAATAGTGCCCCTCACTTTGTCGCTATTACAGCTCAATTCCTAAC
15 TTATAGTCTCTTTTTCCAGTTTGCAGATGCTTATGCAGCTCCTGTACAGGGG
ATTTTACGAGGCTATAAGGATAACAACAAAACCATTTATGATCGGTGCGGG
CTCTTATTGGTTATGTGCTTTGCCATTAGCGGTTATCTTAGAAAAAATAG
CCAGTTAGGTCCGTTTGCCTATTGGATTGGTTTAATCACAGGTATTTTTGTT
TGTGGTCTATTTCTAAACCAACGTCTGCAAAAGATTAAGAAGTTGTATTAT
20 TAA

(SEQ ID NO: 60)

MYQTQTNKEKFVLFLKLFIPVLIYQFANFSATFIDSVMTGQYSQLHLAGVSTAS
NLWTPFFALLVGMISALVPVVGQHLGRGNKEQIRTEFHQFLYLGLLSLILFLI
25 MQFIAQPVLGSLGLEDEVLA VGRGYLNYMLIGIMPLVLFSICRSFFDALGLTRL
SMYLMLLILPFNSFFNYMLIYGKFGMPRLGGAGAGLGTSPTYWAIFIGIIIVMS
LHPQIKTYHIWTLERIKAPLIIEDIRLGLPIGLQIFAEVAIFAVVGLFMAKFSSIIA
AHQAAMNFSSLMYAFPLSISTALAITISFEVGAERFQDATTYSRIGRLTAVGITS
GTLFLFLFRENVAA MYNSAPHFVAITAQFLTYSLFFQFADAYAAPVQGILRG
30 YKDTTKPFMIGAGSYWLCALPLAVILEKNSQLGPFA YWIGLITGIFVCGLFLNQ
RLQKIKKLYY*

Sequence description:

35 A] Length: 1341 bp - 447 aa (full length gene)
B] Shine-Dalgarno sequence present upstream of
ATG start codon, There is a potential signal
peptide sequence

40

ID-106

Clone 2-4A

(SEQ ID NO: 61)

5 TTGCTAGTTTCTTCTCTAGTTTCTTGTTCAATTTTTCTTGTCATTTTCGTCGTT
GTCTTCATCAACACGAAATAAGTCTATAAACTTATCAAATAATTTCATAGA
CTTATTATATCAATTTTCAATAAAAATGCTATAATAAAACCATGTCATTTTCA
TTAAAAATTAGAAATCCATACGGTGAACATACCGTTAAAGAACTCCTTGA
AGATTATTTTTTGATTCCACGTAAGATTAGACATTTTTTGCGTGTTAAAAAA
CATGTACTTATAAACAATGAATTCATTAATTGGCAAACCTGTCGTCCTCAAGAA
10 AACGATACTATTACCTTAATCTTTGATGATGAGGATTACCCTACTAAAAAA
ATTCCTCTGGGCAGAGCAGAGCTTATTGATTGTCTTTATGAGGATGAACAT
CTTATTATCGTTAATAAACCTGAAGGTATGAAAACCTCACGGTAACCAACCA
AATGAAATAGCACTGTAAATCATGTATCTGCCTATTCTGGACAAACATGC
TATGTTGTTTCATCGCCTAGATATGGAGACCAGTGGAGCTGTTTTATTTGCT
15 AAAAATCCATTTATACTTCCCCTTATCAATCAACGCTTAGAACGAAAAGAA
ATTTGGCGTGAATATTGGGCTTTAGTTGAAGGAAAATTTTCACCTAAGCAT
CAAGTTTTGAGAGACAAAATTGGACGGAACCGTCATGACAGACGTAAACG
AATCATTGATTCTAAAAACGGTCAACATGCTATGACAATCATTGACGTTTT
GAAGTATATCCAAAATAGTAGTCTCATAAAATGCCGACTGGAAACCGGAA
20 GAACCCATCAAATTCGCATTCACTTATCTCATCACGGACATCCTTTAATAG
GAGATCCCCTCTACAACCCTTCTTCTAATAATGAAAGGTTAATGCTACACG
CTCACCGATTGACTCTATCCCATCCATTAACCTTGCGAAACTATTAGCGTAG
AGGCCCTTCATCTACTTTTCGAGAAGGTTTTAAACAATTATAAAAAAAGGAG
TTGGATAA

25

(SEQ ID NO: 62)

MLVSSLVSCSFVLVISSLSSSTRNKSINLSNNFIDLLYQFSIKCYNKTMSFSLKIR
NPYGEHTVKELLEDYFLIPRKIRHFLRVKKHVLINNEFINWQTVVQENDTITLIF
DDEDYPTKKIPLGRAELIDCLYEDEHLIIVNKPEGMKTHGNQPNEIALLNHVSA
30 YSGQTCYVVHRLDMETSGAVLFAKNPFIPLINQRLERKEIWREYWALVEGKF
SPKHQVLRDKIGRNRHRRRKRIIDSKNGQHAMTIDVLKYIQNSSLIKCRLETG
RTHQIRIHLSHHGHPLIGDPLYNPSSNNERLMLHAHRLTLSHPLTCETISVEAPS
STFEKVLNNYKKGVG*

35

Sequence description:

- 40 A] Length: 1029 bp - 343 aa (Full length gene sequence)
B] No obvious Shine-Dalgarno sequence upstream
of the putative TTG start codon. Possesses a
potential leader peptide sequence.

ID-107

Clone 2-54

5 (SEQ ID NO: 63)
GAACTAAATGCAACTCAACCTAATAATAGAACTACCTATATTATACCCGAA
AGCAGTCATTCCATTGCAGAACAACAGAGATTCCTGATAGAATCAAAGGG
TTCTTCGGTTGCATTACTTAATAGCGATGAATTTAGAAAGACAGCGGGAGA
GGATAGAGGTTTTTGAAAGGGATAAGTTGAGGTCTTTGGATATCATTCTTAA
10 GGGAGATTTATCGACAAGTAATGTCATAGGTAATACGGACATTGCTAGTC
AGATATCGTTGGGCTTTAAAAAGAATGCGATGCAGGAACACCATCTTACT
AAAACATTCTCTCAAAAGGATGGAAAGTTATCGTCTGTTATAGAGGGGGAT
GCTTGCTATTGGCAAAGAGAGAAAGTAGAGAGAAAGAAATAAAATATAGTGGTA
ATTTATGGCAAAAATTAAAAGCTAAGGCACACTGCCTTGTTTGCTGTGTTG
15 ATAATTTGAATTTTGAAGATATAAAATCTTATTTTCAATATTATTGTCATCT
AAACCATCAGCTCAAATTACCTAAAGGTGCTATACTTTCTGCTAAAACAGA
AGTATATAGGGGAGGAGATTTTGGGAGAAAAAATAAAGATAATGTGTTTG
GTTACCGTATCCCCTCATTATTGAAAACCCAAAAAGGAACTTTACTTGCGG
GAGCTGATGAAAGAATTGAGCAAGCTTGTGATTGGGGAAACATAGGAATG
20 GTTATTCGCCGTAGTGAGGATGATGGTGTCACTTGGGGAAAAAGAGAAAC
TATTGTCAATCTCCGTAATAACCCTAGAGTTCCGCTAGTTACTAGTGGTGA
CTATAGTGGCTCACCTATTAATATGGATATGGCATTAGTTCAAGATACTAG
CTCCAAGACGAAACGTATTTTTTCAATATATGATATGTTTCCAGAAGGAAG
AGGCGTTATTAGTATTGCTAACACACCTGAAAAAGAATATACCCAAATCG
25 GAGGACAGTCTTATCTTAATTTATATAATAATGGAAAGAAATCGAAGGTTT
TACTATCCGTGACAAAGGTATTGTATATAATTTTAAAGGGAAAAAGACTG
ATTATCATGTTATAACAGAACTACTAAAAGTGACCATTCAAATCTAGGGG
ATATTTATAAGGGAAAAACAGCTACTTGGAATATATATTTTACAAACATA
AAACGTCACCATTTTCGTTTAGCAAAATCAAGCTATGTGTGGATGTCATATA
30 GCGATGATGATGGTAGGACATGGTCATCACCTAGAGATATAACAGCAAGT
CTTCGTCAGAAAGGCATGAAATTTTTTGGGAATAGGACCTGGAAAAGGTAT
AGTTTTTAAATGGGGGGCCACACGCTGGTCGTATTATTATTCCTGCCTATTCT
ACGAATTGGAAATCTCATCTAAGAGGTTTACAATCTTCACGCCTAATTTAT
TCAGACGACCATGGAAAAACGTGGCATACTGGAAAAGCAGTTAATGATAA
35 CCGTATACTTTCTAATGGTGAAAAAATTCACCTCCTTAACAATGGATAATAA
AAAAGAACAAAATACAGAATCCGTACCCGTTCAATTGAAAAATGGGGACA
TTAAGTTATTTATGAGGAATCTAACTGGTAACCTAGAAGTAGCCACAAGTA
AAGACGGCGGGGAGACTTGGCAAACCATGTAAACGATATAAGGAAATT
CATGATGCTTACGTCCAACCTATCAGCTATTCGCTTTGAGCATGACAAAAAA
40 GAGTATATTTTATTAGTGAATGCTAATGGGGCCAGGGAAGAAGTGCCAAGA
TGGATATGCACGTCTAGCGCAAGTTAATCGAAATGGTAGTTTTAAGTGGTT
ATATCACCATCACATTCAAGATGGTTCGTTTGCTTACAACCTCTGTTCAACA
ACTTAATAATGATCAATTTGGTGTCTTTATGAACATAGAGAAAAACATCA
AAATAGTTTTACTTTAAATTACAAAGTTTTTAATTGGAGTTTTCTTAGTCAA

AATACAGAGAAGCAAGGCACTTTATGGGAGAAAATGGCAGCAAATTGGCA
TGTTTTGTTTAAATTTTATTTATGA

(SEQ ID NO: 64)

5 ELNATQPNNRTTYIIPESHSHIAEQQRFLIESKGSSVALLNSDEFKRTAGEDRGF
ERDKLRSLDIIPKGDLSNSVIGNTDIASQISLGFKKNAMQEHLTKTFSQKDG
KLSSVIEGMLAIGKEKVEKEIKYSGNLWQKLKAKAHCLVCCVDNLFEDIKS
YFQYYCHLNHQLKLPKGAILS AKTEVYRGGDFGRKNKDNVFGYRIPSLLKTQ
KGTLLAGADERIEQACDWGNIGMVIRSEDDGVTWVGKRETIVNLRNNPRVPL
10 VTSGDYSGSPINMDMALVQDTSSSKTKRIFSIYDMFPEGRGVISIAN TPEKEYTQI
GGQSYLNLYNNGKKS KVFTIRDKGIVYNFKGKKTDYHVITETTKSDHSNLGDI
YKGKQLLGNIYFTKHKTSPFRLAKSSYVWMSYSDDDGRTWSSPRDITASLRQ
KGMKFLGIGPGKGIVLKWGPHAGRIIPAYSTNWKSHLRGSQSSRLIYSDDHG
KWTHTGKAVNDNRILSNGEKIHSLTMDNKKEQNTESVPVQLKNGDIKLFMRN
15 LTGNLEVATSKDGGGETWQNHVKRYKEIHDA YVQLSAIRFEHDKKEYILLVNA
NGPGKKCQDGYARLAQVNRNGSFKWLYHHHIQDGSFAYNSVQQLNNDQFG
VLYEHREKHQNSFTLNYKVFNWSFLSQNTEKQGT LW EKMAANWHVLFKFYL
*

20 Sequence description:

- A) Length: 2052 bp - 684 aa (partial gene sequence)
- B) N-terminus has yet to be determined

25

ID-108

Clone 2-61

30

(SEQ ID NO: 65)

ATGCCTAAATTAATCGTATCTTTCCTCTGCATTTTATTATCCCTGACTTGTG
TAAACTCTGTGCAAGCTGAAGAACATAAAGATATTATGCAAATTACCCGA
GAAGCCGGATATGATGTTAAAGATATTAATAAACCTAAAGCGTCTATCGTT
ATTGACAATAAAGGTCATATTTTGTGGGAAGATAACGCCGATTTAGAACGT
35 GATCCCGCTAGCATGTCTAAAATGTTTACTTTATATTTACTATTTGAAGACT
TAGCTAAAGGAAAAACAAACCTCAACACCACAGTGACTGCAACAGAAACA
GACCAAGCCATAAGTAAGATTTATGAAATTAGTAATAACAATATTCATGCT
GGGGTTGCTTATCCTATTCGTGAACTGATTACTATGACGGCTGTCCCGTCA
TCTAATGTAGCAACTATTATGATTGCTAACCACTTATCACAAAACAATCCT
40 GACGCCTTTATTAAACGAATCAATGAAACCGCCAAGAACTCGGTATGAC
AAAAACTCACTTTTATAACCCCAGTGGGGCGGTAGCGAGTGCTTTTAATGG
ACTTTACTCCCCAAAAGAATACGATAACAATGCTACTAACGTTACGACTGC
ACGTGATCTATCAATTTTAACCTATCATTTTCCTTAAAAAATACCCTGATATA
CTGAAC TATACAAAATATCCTGAAGTCAAGGCCATGGTTCGGAAC TCCTTAT

GAAGAAACATTTACAACCTTATAACTACTCTACCCCCGGCGCTAAATTTGGA
TTAGAAGGAGTAGATGGCTTAAAAACTGGTTCTAGCCCTAGCGCTGCTTTT
AATGCCTTAGTTACAGCTAAACGCCAGAATACTCGCTTGATAACTGTGGTT
TTAGGAGTTGGCGATTGGTCAGACCAAGACGGAGAGTACTATCGTCATCC
5 GTTTGTCAACGCTCTTGTAGAAAAAGGTTTTAAAGACGCTAAAAATATTTC
TTCTAAAACTCCTGTATTAAAAAGCCGTTAAACCTAAAAAAGAAGTTACTAA
AACCAAAACTAAATCTATTCAAGAACAGCCTCAAACAAAAGAAGTGGT
GGACAAAAACAGATCAATTTATCCAATCACATTTTGTATCTATTTTAATTG
TTCTGGGCACCATCGCTAGCCTTTGTCTTTTAGCTGGGATAGTATTACTTAT
10 AAAGCGCTCTAGATAA

(SEQ ID NO: 66)
MPKLIVSFLCILLSLTCVNSVQAEHKDIMQITREAGYDVKDINKPKASIVIDN
KGHILWEDNADLERDPASMSKMFTLYLLFEDLAKGKTNLNTTVTATETDQAI
15 SKIYEISNNNIHAGVAYPIRELITMTAVPSSNVATIMIANHLSQNNPDAFIKRINE
TAKKLGMTKTHFYNPSGAVASAFNGLYSPKEYDNNATNVTTARDLSILTYHF
LKKYPDILNYTKYPEVKAMVGTPYEETFTTYNYSTPGAKFGLEGVDGLKTGS
SPSAAFNALVTAKRQNTRLITVVLGVGDWSDQDGEYYRHPFVNALVEKGFK
DAKNISSKTPVLKAVKPKKEVTKTCTKSIQEQPQTKEQWWTKTDQFIQSHFVS
20 ILIVLGTIASLCLLAGIVLLIKRSR*

Sequence description:

25 A) Length: 1188 bp - 396 aa (full length gene)
B) Shine Dalgarno sequence present upstream of
ATG start codon, possesses a potential signal
peptide

30 ID-109

Clone 45

35 (SEQ ID NO: 67)
ATGACTGAAAAATATTATAATTGGGCAACGCTTGGAACCGGCGTTATTGCC
AACGAATTAGCCCAAGCACTGGAAGCACGTGGACAAAAATTATATTCTGT
AGCTAATAGAACTTACGACAAAGGACTTGAATTTGCTAACAAATATGGTA
TCCAAAAAGTTTATGATCACATAGATCAAGTATTTGAAGAEECTGAAGTGG
40 ATATCATTTTATATCTCTACTCCCCACAATACTCACATCTCATTTTTTACGAAA
GGCTTTAGCAAATGGTAAGCACGTTCTTTGCGAAAAATCTATTACTTTAAA
TAGTACTGAGCTTAAAGAAGCCATAGATTTAGCCGAAACTAACCATGTTGT
CTTAGCTGAAGCCATGACTATTTTTCATATGCCAATTTACCGCCAATTAAA
AACATTAGTTGATAGTGGAATAATTAGGACCGTTAAAAATGATTCAAATGA

ATTTCGGAAGTTATAAAGAATATGATATGACTAACCGTTTTTTCAGTCGTG
ACCTAGCAGGCGGTGCTTTGCTGGACATTGGTGTTTATGCACTTTCTTGTAT
TCGCTGGTTTATGTCAGAAGCACCTCACAACATTACCTCTCAAGTTACATT
TGCACCAACAGGGGTTGATGAACAAGTTGGTATCCTACTAACCAACCCAG
5 CAAATGAGATGGCGACTGTCAGCCTTAGTTTACATGCAAAACAACCTAAA
CGAGCAACTATCGCTTACGATAAAGGCTACATTGAACTTTTTGAATATCCG
CGAGGACAAAAGGCAGTTATTACTTATACTGAGGATGGGCATCAAGATAT
TATCGAAGCTGGCAAAACTGAAAATGCTCTCCAATATGAGGTAGCTGATA
TGGGAAGAAGCCATTTTCAGGAAAAAATAACCACATGTACTTAACTATAACC
10 AAAGATGTTATGGATATCATGACACAGCTACGTCAAGAATGGGGATTAC
CTACCCAGAAGAAGAAAAATGA

(SEQ ID NO: 68)
MTEKYYNWATLGTGVIANELAQALEARGQKLYSVANRTYDKGLEFANKYGI
15 QKVYDHIDQVFEDPEVDIIYISTPHNTHISFLRKALANGKHVLCEKSITLNSTEL
KEAIDLAETNHVVLAEAMTIFHMPYRQLKTLVDSGKLGPLKMIQMNFGSYK
EYDMTNRFFSRDLAGGALLDIGVYALSCIRWFMSEAPHNITSQVTFAPTGVDE
QVGILLTNPANEMATVSLSLHAKQPKRATIAYDKGYIELFEYPRGQKAVITYT
EDGHQDIIIEAGKTENALQYEVADMEEAISGKTNHMYLNYTKDVMDIMTQLR
20 QEWGFTYPEEEK*

Sequence description:

25 A] Length: 984 bp - 328 aa (full length gene)
B] Shine Dalgarno sequence present upstream of
ATG start codon, possesses a potential signal
peptide

30 ID-110

Clone 2-2

35 (SEQ ID NO: 69)
GTGTATTCTCCTGTAAATCTTCTAAAGGAAAAGTGATATTGTTAAAAAGT
GATTTTCTAAAGAGCTTCATAGAAAGGAGAGGAAATATTTGTTTT

40 (SEQ ID NO: 70)
MYSPVKSSKGKVILLKSDFLKSFIERRGNICF

Sequence description:



A] Length: 96 bp - 32 aa (partial sequence)
B] GTG start codon - no obvious Shine-Dalgarno
sequence
Possesses a potential signal peptide

5

ID-111

Clone 2-3

10

(SEQ ID NO: 71)

15

AAATACTGTATCATTGCAACCTCAAATGCAGGTTTTGGAAACGAAGCATTT
ACAGGTGACAGCGATAAAGACTTGAAAATTATGGAACGAATTTCTCCATA
TTTCCGTCCAGAATTTCTAAATCGTTTCAATGGTGTTATTGAATTCTCTCAC
CTAAGCAAAGATGACTTAAGCGAAATTGTAGATTTGATGCTTGATGAAGTT
AACCAAACAATTGGCAAAAAAGGAATTGACCTTGTGGTAGATGAAAATGT
TAAATCACACTTAATTGAACTGGGTTATGACGAAGCAATGGGAGTACGTC
CATTGCGCCGTGTCATCGAGCAAGAAATTCGAGATCGCATCACAGACTACT
ATCTCGATCATAACAGACGTTAAACACCTAAAAGCTAATTTGCAAGATGGCC
AAATCGTCATTTCTGAAAGATAA

20

(SEQ ID NO: 72)

25

KYCIATSNAGFGNEAFTGDSKDLKIMERISPYFRPEFLNRFNGVIEFSHLSD
DLSEIVDLMLDEVNQTIGKKGIDLVDENVKSHLIELGYDEAMGVRPLRRVIE
QEIRDRITDYYLDHTDVKHLKANLQDGQIVISER*

Sequence description:

30

A] Length: 429 bp - 143 aa (partial sequence)
B] N-terminus yet to be elucidated. This gene
was not in frame with nuc

35

ID-112

Clone 2-5

40

(SEQ ID NO: 73)

ATGTCAATGAATTTTTCATTTTACCACAATATTGGTCCTATTTTAATTATG
GTGTGATGGTAACCATTATGATTTCACATGTGTTGTTTTTTTGGAACTAT
TATAGGCGTGTTAATTGCTTTAGTAAAGCGTACTAATTTACATTTTCTCACA
ATATTAGCTAATTTCTATGTATGGGTATTTCGTGGGACACCGATGGTAGTT
CAAATTATGATTGCTTTCGCATGGATGCATTTTAACAATTTACCAACAATT

5 AGCTTTGGTGTTTTAGATTTAGATTTTACACGACTTTTACCTGGTATCATTA
TCATTTTCCTTAAATAGTGGTGCCTATATTTTCGGAAATTGTACGTGCAGGGA
TTGAGGCTGTACCATCTGGACAAATAGAAGCAGCTTACTCGTTGGGGATTC
GACCTAAAAATACACTTCGCTATGTTATCTTACCCCAAGCTTTTAAAAATA
TTTTACCTGCTCTAGGGAATGAATTTATTACAATTATTAAAGATAGTGCTCT
CCTTCAAACCTATTGGTGTGCATGGAATTATGGAACGGAGCACAAATCAGTTGT
AACGGCTACTTACTCACCAGTTGCACCGTTATTATTTGCAGCATTTTACTAT
TTAATGTTGACAACGATTCTCTCAGCTTTGTTAAAAACAAATGGAGAAATAT
CTTGGGAAAGGGGTAAAAATAGATGGTTGA

10

(SEQ ID NO: 74)

MSMNFSFLPQYWSYFNYGVMVTIMISTCVVFFGTIIGVLIALVKRTNLHFLTIL
ANFYVWVFRGTPMVVQIMIAFAWMHFNNLPTISFGVLDLDFTRLLPGIIIISLNS
GAYISEIVRAGIEAVPSGQIEAAYSLGIRPKNTLRYVILPQAFKNILPALGNEFTI
15 IKDSALLQTIGVMELWNGAQSVVTATYSPVAPLLFAAFYYLMLTTILSALLKQ
MEKYLKGKGVKIDG*

20 Sequence description:

- A] Length: 699 bp - 233 aa (full length gene)
B] Shine-Dalgarno sequence preceded the 'ATG'
start codon. Possesses a potential leader peptide
sequence.

25

ID-113

Clone 2-7

30

(SEQ ID NO: 75)

ATGAAAGACCTATTACGAAATAGTCTAGAGCAAAGTGGAAATTTAAGTTT
TCAAGATATGATTTTACATATTCTTGTAGCAGCTTTATTGAGTGTAGTTATT
TATGTTTCCTATGCTTATACGCATAGTGGAACCTGCCTATAGTAAAAAGTTT
35 AATGTTTCATTAATGACATTGACGGTCTTGACTGCAACAGTAATGACCGTT
ATTGGTAATAATGTAGCCTTGTCATTGGGTATGGTCGGTGCCTTGTCAGTT
GTTCGTTTTAGGACAGCCATAAAAGATTCAAGAGATACAGTTTATATTTTT
TGGACCATAGTTGTTGGTATCTGTTGTGGTGTCTGGTGACTATGTGGTAGCT
GCATTAGGAAGTAGCGTTATCTTTATCTTATTATGGGTATGGGACGTGTT
40 AAAAACGAGAATCGTATGTTATTGATTGTGAAGTGCGATAGAACACTAGA
AGTTGATTTAGAAGGAATTTTCTTCCAATATTTTGACGGAAAAGCTGTTCA
GCGTGTTAAAAATTCAACAATACTATTGAAATGATTTTCGAAATCTC
TAGAAAAGATTACGATAAGCAACTCCATGTAGATAATCAGTTAACTGAAA

AAGTGTACCAATTGGGAAATATTGATTATTTCAACATTGTTAGCCAAAGCG
ACGAAATCAATGGGTAG

(SEQ ID NO: 76)

5 MKDLLRNSLEQSGNLSFQDMILHILVAALLSVVIYVSYAYTHSGTAYSKKFNV
SLMTLTVLTATVMTVIGNNVALSLGMVGALSVVRFRTAIKDSRDTVYIFWTIV
VGICCGVGDYVVAALGSSVIFILLWVMGRVKNENRMLLIVKCDRTLEV DLEGI
FFQYFDGKAVQQRVKNSTTNTIEMIFEISRKDYDKQLHVDNQLTEKVYQLGNID
YFNIVSQSDEING*

10

Sequence description:

15 A] Length: 678 bp - 226 aa (full-length gene)
B] ATG start codon is preceded by a Shine-
Dalgarno sequence-Possesses a potential leader
peptide sequence

20 ID-114

Clone 2-8

(SEQ ID NO: 77)

25 AAAAATTCATTTTAGATTCATTTTACGACTATATACTCAGAAGTACCAAAC
CTAATCCAAGGTTTGAAAAAAGAAAGAAGGAAGTCAGTATGACAAACTAT
AAAAACAAC TTTAAAGATGAGGCTATACGTGTTGAAGAGACAACAAAAGA
ATCATTTTACGATGTTGATATTGCCTTGTTTTTCAGCTGGTGGATCTATTTCA
GCAAAGTTCGCTCCTTATGCAGTAAAGTCTGGAGCAGTTGTAGTAGATAAC
30 ACGTCATATTTTCGTCAGAATCCTGATGTTCCACTAGTTGTTCCCTGAAGTAA
ATGCTCATGCCATGATTGGTCATAATGGTATCATAGCTTGTCCCAATTGTTC
TACTATTCAAATGATGATTGCTTTAGAGCCCATTCGTCAAAAATGGGGGGAT
AGAGCGTGTTATAGTTTCCACCTATCAAGCTGTTTCGGGGTTCAGGTGCACG
TGCTGTTGAAGAACTAAGGAACAGTTGAGACAAGTTTT

35

(SEQ ID NO: 78)

KFILD SFYDYILRSTKPNPRFEKRKKEVSMTNYKNNFKDEAIRVEETTKESFYD
VDIALFSAGGSISAKFAPYAVKSGAVVVDNTSYFRQNPDVPLVVPEVNAHAMI
GHNGI IACPNCSTIQMMIALEPIRQKWGIERVIVSTYQAVSGSGARAVEETKEQ
40 LRQV

Sequence description:

A] Length: 499 bp - 165 aa (partial sequence)
B] N-terminus has yet to be determined

5 ID-115

Clone 2-9

(SEQ ID NO: 79)

10 ATGACAAATGAATTGATAATGCAAGCTTTTGAGTGGTATTTACCTAGTGAT
GGGAATCACTGGAAGAAATTAGAGGAGTCTATATCAGACCTTAAAAAACT
TGGAATTAGTAAAATCTGGTTACCACCAGCATTTAAGGGAACTAGCAGTG
ATGATGTAGGATATGGTGTTTATGATCTCTTTGATTTAGGAGAATTTGACC
15 AGAATGGAACAATTAGAACAAAATATGGTAGGAAAGAAGAGTATCTAAA
GCTTATTAAGTCGTTAAAGGCAAATGGCATTAAACCGTTTGCAGATATCGT
TCTTAACCATAAAGCCAATGGTGATCATAAAGAAAAATTTCAAGTCATCA
AAGTCAATCCTGAAAATCGTCAAGAAGCATTAAAGTGAACCCTATGAGATT
GAAGGATGGACGGGATTTGATTTCCCAGGTAGACAGGGTGAGTACAATGA
TTTT

20

(SEQ ID NO: 80)

MTNELIMQAFEWYLPDGNHWKKLEESISDLKKLGISKIWLPPAFKGTSSDDV
GYGVYDLFDLGEFDQNGTIRTKYGRKEEYLKLIKSLKANGIKPFADIVLNHKA
25 NGDHKEKFQVIKVNPNRQEALSEPYEIEGWTGFDFPGRQGEYNDF

25

Sequence description:

30 A] Length: 456 bp - 152 aa (partial sequence)
B] ATG start codon is preceded by a Shine-
Dalgarno sequence, no leader peptide sequence.

35 ID-116

Clone 2-10

(SEQ ID NO: 81)

40 ATGGAGGTTCTTATGAAGAAAGTGTTAGTAAGTAGTCTTTTGGTTTTAGGG
ATTACGATAACGTTACAACCAGTAGTTGAGGCTAAGGGGGCCAAAAGTAGC
TTATACACAAGAGGGAATGACTGCTCTTTCGGACACAAATAAAGATAAAG
TCACTACTATTTCTATTGACGAGATTCAAAAAAGCTTAGAAGGTAAGAAGC
CGATTACTGTTAGTTTTGATATTGATGATACACTGCTTTTCAGTAGTCAATA

TTTTCAATATGGTAAAGAATATGTAACCTCCTGGATCGTTTGATTTTCTTCAT
AAACAAAAATTCTGGGATCTTGTTGCAAAACGAGGAGATCAAGATTCCAT
TCCCAAAGAATATGCTAAAAAATTAATTGCTATGCATCAAAAACGAGGAG
ATAAAATTGTTTTTATAACAGGTAGGACAAGAGGGGTCAATGTATAAGGAG
5 GGCGAGGTTGATAAAACAGCTAAAGCCTTAGCTAAAGATTTTAAATTTGTA
CCATCTGAT

(SEQ ID NO: 82)

MEVLMKKVLVSSLLVLGITITLQPVVEAKGPKVAYTQEGMTALSDTNKDKVT
10 TISIDEIQKSLEGKKPITVSFDIDDTLLFSSQYFQYGKEYVTPGSFDLHKQKFW
DLVAKRGDQDSIPKEYAKKLIAMHQKRGDKIVFITGRTRGSMYKEGEVDKTA
KALAKDFKFVPSD

15 Sequence description:

A] Length: 516 bp - 172 aa (partial sequence)
B] ATG start codon is preceded by a Shine-
20 Dalgarno sequence, Possesses a leader peptide
sequence.

ID-117

25 Clone 2-17

(SEQ ID NO: 83)

ATGCTTAAAAGATTATTTACTGAAGATGGGGAATTGACAAAGATTAGTCGT
CGTTTCGTTTGGATGTTAGTGGTTATCTATTGTCTTATTATTGTCAGGATGT
30 GTTTTGGGCCTCAAATTATGATTGAGGGGGGTATCAACTCCGAATGTTTCAGC
GCTTCGGAAGAATTGTAGCTCTTTTAGTACCATTTAATTCTTTTCGTAGTTT
AGATCAGCTAACTAGCTTTAAAGAGATTCTTTGGGTTATTGGTCAAAAATGT
AGTGAATATTTTACTGCTGTTTCCTCTCATTATAGGGTTACTATCCCTAAAG
CCAAGTTTACGGAAATATAAAAGCGTTATATTACTTGCTTTCTTGATGTCTC
35 TTTTCATAGAGTGTACTCAAGTTGTTTTAGATATTTTAATAGATGCTAATCG
GGTTTTTGAAATCGACGATCTATGGACAAATACCTTAGGCGGTCCTTTTCGC
CCTATGGAGTTATCGAAACATAAAAGGTTGGCTTCTAACTATTAGAAAATG
A

40 (SEQ ID NO: 84)

MLKRLFTEDGELTKISRRFVWMLVVIYCLIIVRMCFGPPQIMIEGVSTPNVQRFG
RIVALLVPFNSFRSLDQLTSFKEILWVIGQNVVNILLFPLIIGLLSLKPSLRKYK
SVILLAFLMSLFIECTQVVLDILIDANRVFEIDDLWTNTLGGPFALWSYRNIKG
WLLTIRK*

Sequence description:

- 5 A] Length: 516 bp - 172 aa (full-length gene)
 B] ATG start codon is preceded by an Shine-
 Dalgarno sequence. Possesses a potential leader
 peptide sequence. C-terminus need further
 confirmation.

10

ID-118

Clone 3-3

15

(SEQ ID NO: 85)

ATGAAAAAGCTTACTTTTATTTGGGATTTAGATGGGACATTAATAGATTTCG
TATGTACCAATTATGGAAGCTCTTGAAGAAACCTATCGTCATTTTGGCTTA
ATATTTGATAAAGAATTAATCCATGAATATATTTTACAGGAATCAGTGGGG
20 CAATTATTGGTAAACCTTTCAGAGGAAGAGCAAATACCTCATGAAAAACT
GAAAGCATATTTTACAAAAGAACAAGAAAGTCGAGATTCTAAAATACATT
TAATGCCATATGCAAAAGAGATTTTAGAATGGACCAAAGAACAAGATATT
CCCAATTTTATGTATACACATAAAGGAGCAAGTACGCATTCAGTGTTGGAA
ACCTTGCAGATCTCTCATTATTTTGATGAAATTTTAAGTGGTGTTCGGGAT
25 TCGAGCGAAAACCACATCCACAAGGGATTAATTATTTAGTTAAACGATATT
CTTTAGATAAATCAATGACTTATTACATAGGAGATCGTCCACTAGATTTGG
AGGTTGCTCAAAATGCTGGTATAAAATCCATAAACTTAAGGTTAGAGAATT
CCAAAGAAAACCTATAATATTTCAAGTCTCAAAGATATAATATCACTTGATT
TCACTCGTTTGGATTAA

30

(SEQ ID NO: 86)

MKKLTFIWDLDGTLIDSYPIMEALEETYRHFGLIFDKELIHEYILQESVGQLL
VNLSEEEQIPHEKLKAYFTKEQESRDSKIHLMPYAKEILEWTKEQDIPNFMYTH
KGASTHSVLETLQISHYFDEILTGVSGFERKPHPQGINYLVKRYSLDKSMTYYI
35 GDRPLDLEVAQNAGIKSINLRLENSKENYNISSLKDIISLDFTRLD*

Sequence description:

- 40 A] Length: 627 bp - 209 aa (Possible Full-length gene)
 B] ATG start codon is preceded by an possible
 Shine-Dalgarno sequence. No obvious leader
 peptide sequence.

ID-119

Clone 3-7

5

(SEQ ID NO: 87)

ATGGAAAAAGAAAAAAATTAGGTCTTTTACCACTAACAATGCTTGTCATT
GGCTCTCTTATCGGTGGCGGAATCTTTGATTTAATGCAAAATATGAGTTCC
AGAGCCGGTTTGGTACCAATGCTTATTGCTTGGGTAATTACTGCTATCGGG
10 ATGGGAACTTTCGTTTTAAGTTTTCAAAATTTATCTGAAAAAAGGCCGGAC
CTAACAGCTGGAATCTTTAGTTACGCTAAAGAGGGGGTTTGGAACTTTATG
GGATTTAAGTCTGCATGGGGTTATTGGTTATCAGCTTGGCTTGGAAATGTT
GCCTACGCTGCACTCTTATTCAGTTCACCTCGGTTATTTCTTTAAATTCTTTG
GTAATGGAAATAATATCATCTCAATTATTGGAGCAAGTATAGTTATTTGGG
15 TTGTCCATTTCTTAATTTTAAGAGGTGTTAATACAGCTGCATTTATTAATAC
CGTAGTTACCTTTGCAAAATTAGTACCTGTTATTATTTTCTTAATTTTCAGCG
TTATTAGCTTTCAAATTTAACATTTTATAGTCTTGATATCTGGGGAAATGGAT
TACATCAATCAATTTTCAACCAAGTCAATTCAACTATGAAAACCGCTGTTT
GGGTATTTATTGGTATTGAGGGCGCCGTTGTCTTCTCAGGTCGTGCTAAAA
20 AACACTCTGATATTGGTAAAGCAAGTATCCTAGCATTATTCACCTATGATTT
CACTTTATGTATTGATTTCTGTTTTATCACTTGGTATCATGTCACGTCCAGA
ACTTGCAAACCTTAAAAACACCAGCTATGGCTTACGTTCTAGAAAAAGCTGT
TGGTCACTGGGGTGCTATCTTAGTTAACCTTGGTGTTATCATTTTCAGTATTT
GGCGCTATTCTTGCTTGGACTTTATTTGCAGCAGAATTACCATATCAAGCT
25 GCTAAAGAAGGTGCTTTTCCTAAATTTTTTGCAAAAGAAAATAAAAACAA
AGCTCCAATCAACTCACTCTTAGTCACTAATCTTTGTGTACAAGCATTCTTA
ATCACGTTCTTATTCACACAAAGTGCTTATCGTTTTGGTTTCGCATTAGCAT
CATCTGCTATCTTAATTCCTTATGCTTTTACAGCACTATATCAATTACAATT
CACACTCCGTGAGGATAAGTCAACTCCAGGACATCAAAAGAATTTAATTA
30 TCGGTATCCTCGCTACAATCTATGCTGTTTACCTTATCTACGCTGGTGGTTT
TGATTACTTACTTTTGACAATGATTGCTTATACTCTAGGTATGATTCTCTAT
ATTAATAATGAGAAAAGATGACAAGCTTGGCGTAATCATGGTCATAGCTGT
TTCCAGTGTGAAATTGTTATCC

35

(SEQ ID NO: 88)

MEKEKKLGLLPLTMLVIGSLIGGGIFDLMQNMSSRAGLVPMLIAWVITAIGMG
TFVLSFQNLSEKRPDLTAGIFS YAKEGFGNFMGFNSAWGYWLSAWLGNVAY
AALLFSSLGYFFKFFGNGNNIISIIGASIVIWVVHFLILRGVNNTAAFINTVVTF
LVPVIIFLISALLAFKFNIFSLDIWGNGLHQSIQVNSTMKTAVWVFIGIEGAV
40 VFSGRAKKHSDIGKASILALFTMISLYVLISVLSLGIMSRPELANLKTPAMAYV
LEKAVGHWGAILVNLGVII SVFGAILAWTLFAAELPYQAAKEGAFPKFFAKEN
KNKAPINSLLVTNLCVQAFLITFLFTQSAYRFGFALASSAILIPYAFTALYQLQF
TLREDKSTPGHQKNLIIGILATIYAVYLIYAGGFDYLLLTMIAYTLGMILYIKMR
KDDKLGVMVIAVSSVKLLS

Sequence description:

5

A] Length: 1356 bp - 452 aa (partial sequence)

B] ATG start codon is preceded by an possible
Shine-Dalgarno sequence. Possesses a potential
leader peptide sequence.

10

ID-120

Clone 3-8

15

(SEQ ID NO: 89)

ATGAAATTTGAAAAACGGCAGGTCTATTATGTTGTCATAACATTTGCTATT
TGCTATGCTATACAGGCTTATTGGGGAGCTGTTTCTAATATTTTAACTACGC
TTCATAAGGCAATATTTCTTTTTTTGATGGGAGCTGGAATTGCCTATATTAT
20 TAATATTGTAATGTCAGTCTATGAGCGATTATATATAAAGCTTTTTTAAAGG
ATCTAGACTATTAATGGCAATCAAGCGTAGTGTTTCTATGATTTTATCCTAT
GCAACTTTTATTGGTTTAATTGTCTGGCTATTTTCAATTGTCATTCCAGATT
TGATTTCTAGTTTGAGTTCTTTATTGGTTATTGATACCGGAGCACTTGCTAA
ATTGGTTAATAATCTCAATGAAAATAAACAAATTTCTGAGGCTTTAAATTA
25 TATGGGAACAGATAAAGACTTAGTTTCTACTTTAAGTGGTTATAGCCAGCA
GATTTTGAAGCAAGTTTATCTGTTTTAACAAATTTACTAACCTCAGTTTCC
TCTATTGCGGCAACACTTCTGAATGTTTTTGTTAGTTTTATTTTTTCAATTTA
CGTTTTGGCAAACAAGGAGCAGTTGGGACGTCAATTTAATTTGTTAATTGA
TACCTATTTAGGTTCAACAGGCAAAACATTCCATTACGTTTCGTCATATCCTT
30 CATCAACGTTTCCATGGTTTTTTTTGTAAGCCAAACTTTAGAAGCTATGATTT
TAGGAAGTTTGACGGTTATTGGTATGTTGATCTTCCAATTTCTTATGCTTT
AACAGTTGGGGTTTTAGTTGCTTTTACAGCTCTAATACCGGTTGTGGGAGC
CTACATTGGTGTTACAATCGGTTTCATCTTAATTGCTACTGAATCGCTTACT
GAAGCATTCTTGTTTGTTCTTTTCTTGATCCTTTTACAACAATTTGAGGGAA
35 ATGTCATTTATCCGAAAGTTGTCGGTGGATCGATTGGACTGCCTTCTATGT
GGGTTTTAATGGCTATTACTATCGGAGGTGCTTTATGGGGGATCTTAGGCA
TGTTACTTGCTGTTCTGTTGCAGCTACTATCTATCAGATTGTAAAAGATCA
TATTATCAAGCGACAAACGCTTAGAAATCGTGCACGAACCTATCGTTAA

40

(SEQ ID NO: 90)

MKFEKRQVYYVVITFAICYAIQAYWGAVSNILTTLHKAIFPFLMGAGIAYIINI
VMSVYERLYIKLFGKSRLMAIKRSVSMILSYATFIGLIVWLFSIVPDLISLSS
LLVIDTGALAKLVNNLNENKQISEALNYMGTDKDLVSTLSGYSQQILKQVLSV
LTNLLTSVSSIAATLLNVFVSFIFSIYVLANKEQLGRQFNLLIDTYLGSTGKTFH

YVRHILHQRFHGFFVSQTLEAMILGSLTVIGMLIFQFPYALTVGVLVAF TALIP
VVGAYIGVTIGFILATESLTEAFLFVLFLILLQQFEGNVIYPKVVGGSIGLPSM
WVLMAITIGGALWGILGMLLAVPVAATIIYQIVKDHIKRQTLRNRARTYR*

5

Sequence description:

- 10 A] Length: 1134 bp - 378 aa (full-length gene)
 B] ATG start codon is preceded by an typical
 Shine-Dalgarno sequence. Possesses a potential
 leader peptide sequence.

15 ID-121

Identical to ID-68, as described in WO 00/06736

20

ID-122

Clone 3-16

25

(SEQ ID NO: 91)

GTGATTACAATTAAAAAGGAATCTGTTATCAAACCTATTGAAGTATGCTTTT
GGCATTATAATGGGATTTATTATCTTAGCTATTGTAATAGGTGGGCTCCTA
TTTGCATACTACGTTAGTCGTTCTCCGAAATTAACCGATCAAGCTTTAAAA
TCCGTTAACTCTAGTTTGGTTTATGATGGTAATAATAAACTTATTGCCGATT
30 TAGGCTCAGAAAAGCGTGAAAGTGTTAGTGCGGATAGCATTCCACTAAAT
TTGGTTAACGCTATCACTTCTATAGAAGATAAACGTTTCTTTAAACATAGA
GGTGTCGATATTTATCGTATTTTAGGTGCAGCTTGGCATAACCTTGTTAGTA
GTAATACGCAAGGTGGTTCAACCCTTGATCAACAGTTGATTAAACTGGCTT
ACTTTTCTACCAATAAATCTGACCAAACGTTAAAACGTAAATCACAGGAA
35 GTTTGGCTTGCGCTTCAAATGGAGCGTAAATACACCAAAGAAGAAATTCTT
ACTTTCTATATTAATAAAAGTTTATATGGGAAATGGGAATTATGGTATGAGA
ACAACAGCTAAATCATACTTTGGTAAAGACCTAAAGGAATTATCTATTGCA
CAACTTGCTTTGCTCGCTGGTATTCCTCAAGCACCTACACAATATGACCCTT
ATAAAAACCCAGAATCTGCTCAAACAAGACGTAATACCGTTCTTCAGCAG
40 ATGTATCAAGATAAAAACATTTCTAAAAAGGAATACGACCAAGCTGTTGC
AACTCCAGTAACTGATGGCTTAAAAGAATTAAAGCAAAAATCTACTTATCC
AAAATATATGGATAACTACTTAAAACAAGTTATTAGTGAAGTTAAACAAA
AAACTGGTAAAGATATCTTTACTGCTGGGCTAAAAGTGTATACTAATATCA
ACACTGATGCACAAAAACAACCTATATGACATCTACAACAGTGATACTTAC

ATCGCTTATCCAAACAATGAATTACAAATAGCATCTACCATCATGGATGCG
ACTAATGGTAAAGTCATTGCACAATTAGGCGGGCGTCATCAGAATGAAAA
TATTTTCAATTTGGGACAAATCAATCTGTCTTAACAGACCGCGATTGGGGTTC
TACAATGAAACCTATCTCAGCTTATGCACCTGCTATTGATAGTGGTGTCTA
5 TAATTCAACAGGTCAATCATTAAACGACTCAGTTTACTACTGGCCTGGTAC
TTCTACTCAACTATATGACTGGGATCGTCAATATATGGGTTGGATGAGTAT
GCAGACCGCTATTCAACAATCACGTAACGTCCCTGCTGTCAGAGCACTTGA
AGCCGCTGGATTAGACGAAGCAAAATCTTTCCTTGAAAAATTAGGCATAT
ACTATCCAGAAATG

10

(SEQ ID NO: 92)

MITIKKESVIKLLKYAFGIIMGFILAIIVIGLLFAYYVSRSPKLTQALKSVNSS
LVYDGNKLIADLGSEKRESVSADSIPLNLVNAITSIEDKRFFKHRGVDIYRILG
AAWHNLVSSNTQGGSTLDQQLIKLAYFSTNKSDQTLKRKSQEVWLALQMER
15 KYTKEEILTFYINKVYMGNGNYGMRTTAKSYFGKDLKELSIAQLALLAGIPQA
PTQYDPYKNPESAQTRRNTVLQQMYQDKNISKKEYDQAVATPVTGDLKELK
QKSTYPKYMDNYLKQVISEVKQKTGKDIFTAGLKVYTNINTDAQKQLYDIYN
SDTYIAYPNNELQIASTIMDATNGKVIAQLGGRHQNENISFGTNQSVLTDRDW
GSTMKPISAYAPAIIDSGVYNSTGQSLNDSVYYWPGTSTQLYDWDRQYMGWM
20 SMQTAIQSRNVPAVRALEAAGLDEAKSFLEKLGIIYPEM

Sequence description:

25 A] Length: 1386 bp - 462 aa (partial sequence)
B] GTG start codon is preceded by an
typical Shine-Dalgarno sequence. Possesses a
potential leader peptide sequence.

30

ID-123

Clone 3-17

35

(SEQ ID NO: 93)

ATGGCTAATGTATATGATTTAGCAAATGAATTAGAACGTGCTGTTCGTGCT
TTACCAGAATACCAAGCAGTTTTAACTGCAAAAGCAGCTATTGAAAATGA
TGCGGATGCACAAGTGCTTTGGCAAGACTTTTTGGCTACCCAATCAAAAGT
TCAAGAAATGATGCAATCTGGCCAAATGCCAAGTCAAGAAGAACAAGATG
40 AAATGTCTAAACTTGGGGGAAAAAATTGAATCCAATGACCTTTTAAAAGTTT
ATTTTGACCAACAACAACGGTTGTCTGTCTATATGTCTGATATCGAAAAAA
TTGTCTTTGCACCCATGCAGGACTTGATGTAA

(SEQ ID NO: 94)

MANVYDLANELERAVRALPEYQAVLTAKAAIENDADAQVLWQDFLATQSK
VQEMMQSGQMPSQEEQDEMSKLGEKIESNDLLKVYFDQQRLSVYMSDIEKI
VFAPMQDLM*

5

Sequence description:

10 A] Length: 336 bp - 112 aa (full length sequence)
B] ATG start codon is preceded by an
typical Shine-Dalgarno sequence. No obvious
potential leader peptide sequence.

15

ID-124

Clone 3-26

20

(SEQ ID NO: 95)

ATGGCAGAAATCACAGCTAAACTTGTAAGAATTGCGTGAAAAATCAGG
TGCAGGCGTTATGGACGCTAAAAAAGCATTAGTAGAAACTGATGGTGACC
TTGATAAAGCGATTGAATTACTTCGCGAAAAAGGTATGGCTAAAGCAGCT
AAAAAAGCAGACCGTGTTGCTGCTGAAGGTTTAACAGGTGTTTATGTTGAT
25 GGTAACGTTGCAGCAGTTATTGAAGTTAA

25

(SEQ ID NO: 96)

MAEITAKLVKELREKSGAGVMDAKKALVETDGDLDKAIELLREKGMAKAAK
KADRVA AEGLTGVYVDGNVA AVIEV

30

Sequence description:

35 A] Length: 230 bp - 76 aa (partial sequence)
B] ATG start codon is preceded by an
typical Shine-Dalgarno sequence. No obvious
potential leader peptide sequence.

40

ID-125

Clone 3-33

(SEQ ID NO: 97)

ATGATAAAAAACCTGTTATTAACAGGTTTTTTTATCATTTAATGACGGAAAA
CTGGACACAAATTATTTTTCTTGTATAATTAAATATATTATTTCTTATCAGG
AGGTTATGATGACATTAGAGAAACGATTAA

5

(SEQ ID NO: 98)

MIKNLLLTGFLSFNDGKLDTNFYFSCHIKYIISYQEVMMTLEKRF

10 Sequence description:

A] Length: 134 bp - 44 aa (partial sequence)

B] ATG start codon is preceded by an
typical Shine-Dalgarno sequence. Possible
potential leader peptide sequence.

15

ID-126

20 Clone 3-41

(SEQ ID NO: 99)

ATGAAAAATAATAAAAAATAATGGTTTTCTGAAAAATTCCTTTATTTACATA
TTATTGATTATTGCGGTTATTACAACCTTTCAATACTATTAA

25

(SEQ ID NO: 100)

MKNNKNNGFLKNSFIYILLIIAVITTFQYYL

30 Sequence description:

A] Length: 94 bp - 31 aa (partial sequence)

B] ATG start codon is preceded by a
possible Shine-Dalgarno sequence. Potential
leader peptide sequence.

35

ID-127

40 Clone 3-42

(SEQ ID NO: 101)

ATGTTAGATATTATCTTATCCGGAATTTGCAAGGATTACTTTGGTCAATTA
TGGCAATTGGCGTGTTTATCACTTTTCGTATCTTAGACATAGCCGATCTCTC

TGCAGAAGGGGCTTTCCCTATGGGGGGCTGCAGTTTGCGCCTTATGTATCGT
TAA

(SEQ ID NO: 102)

5 MLDIILSGISQGLLWSIMAIGVFITFRILDIADLSAEGAFPMGAAVCAALCIV

Sequence description:

10 A] Length: 158 bp - 52 aa (partial sequence)
B] ATG start codon is preceded by a
possible Shine-Dalgarno sequence. Potential
leader peptide sequence.

15 ID-128

Clone 3-43

20 (SEQ ID NO: 103)

ATGGAAATGCCTAAAAGAAATGAATTACTCAATAAAGAAATTAAAATGAG
TATTGATAAACTTAGATATAAAGAACCAGAGAGTGAACATGACAAGCGAC
CTACTTTTTATTTGGTAGTACTTATACTTGTTACTGTAGCAGTTATATTGTC
GTTATTAA

25 (SEQ ID NO: 104)

MEMPKRNELLNKEIKMSIDKLR YKEPESEHDKRPTFYL VVLILVT VAVILSLF

30 Sequence description:

35 A] Length: 161 bp - 53 aa (full-length gene)
B] ATG start codon is preceded by a
possible Shine-Dalgarno sequence. Potential
leader peptide sequence.

ID-129

40 Clone 3-44

(SEQ ID NO: 105)

GTGGTAAGTAAATTGAGTTTAACAACGATTTTTGCATTGCTATTTTCATCA
ATGCTAATTTACGCAACACCTCTTATCTTTACAAGTATTGGGGGAACCTTC

TCTGAACGTGGTGGTATCGTCAACGTTGGTTTAGAAGGAATTATGGTAATT
GGAGCTTTCTCAGGCGTTGTATTAA

(SEQ ID NO: 106)

5 MVSKLSLTTIFALLFSSMLIYATPLIFTSIGGTFSEKGGIVNVGLEGIMVIGAFSG
VVF

Sequence description:

10

A] Length: 179 bp - 59 aa (partial sequence)
B] GTG start codon is preceded by a
possible Shine-Dalgarno sequence. Potential
leader peptide sequence.

15

ID-130

Clone 3-46/47

20

(SEQ ID NO: 107)

ATGAGAATTATTGCAATAACTGAAAAGGTTATAAAAGAACTGTTTCGTGAT
AAAAGAACACTTGCTATGATGTTTTTAGCACCTATTTTAATTATGTTTTTGA
TGAATGTTATGTTTTCTGCGAATAGTAATACAAAAGTTAAGATTGGAACTA
25 TTAACGTTAACACGAAGGTCGTTTCAAATTTAGATAATATTAAGCATATTC
AAGTGAGATCATTTAAATTTAACTCATCTGCTAAAAAAGCACTCAAATCAA
ATAAAATTGATGCTCTTATTTCGGAGGACAATAAATCTTATACTGTCTTCT
ATGCGAATACAGATTCTTCAAAGACGACTTTAACAAGACAAGCTTTTAAA
ACCGCTGTTAATAACAATGAACAGTAAGGAAGTATTTCGCAAGTTAAAATT
30 TTAGCTAATAAGAATCCGAACTAGCACAACTCCTTACAACTCGCTCCAAA
TATATCAAAGAAAAATATAATTACGGAAATAAAAATACAGGCTTTTTTGC
AAAAATGATACCAATACTAATGGGATTTATGGTCTTCTTCTTGGTTTTT

(SEQ ID NO: 108)

35 MRIIAITEKVIKELFRDKRTLAMMFLAPILIMFLMNVMSANSNTKVKIGTINV
NTKVVSNDNIKHQVRSFKFNSSAKKALKSNKIDALISEDNKS YTVFYANTDS
SKTTLTRQAFKTA VNTMNSKELISQVKILANKNP KLAQSLQTRSKYIKEK YNY
GNKNTGFFAKMIPILMGFMVFFLVF

40

Sequence description:

A] Length: 558 bp - 186 aa (partial sequence)
B] ATG start codon is preceded by a

possible Shine-Dalgarno sequence. Potential leader peptide sequence. C-terminus has yet to be determined.

5 ID-131
Clone 3-48

10 (SEQ ID NO: 109)
GTGATTATCGTTATGAGTAAACATCAAGAAATTTTGGAGTACCTAGAAAAT
TTAGCTGTTGGTAAGAGGGTTAGTGTACGCAGTATTTCAAATCATTTAA

(SEQ ID NO: 110)

15 MIIVMSKHQEILEYLENLAVGKRVSVRISNHL

Sequence description:

20 A] Length: 100 bp - 33 aa (partial sequence)
B] GTG start codon is not preceded by a
obvious Shine-Dalgarno sequence. No obvious
leader peptide sequence.

25 ID-132
Clone 2-c53

30 (SEQ ID NO: 111)
ATGTATAGAGAAATTACCGCTGTCGAACACGATCGCTTTGTGAGCGAATCC
AACCAAACAAACCTACTTCAATCTCTTAATTGGCCCAAAGTAAAAGACAA
CTGGGGTAGTCAATTACTTGGCTTTTTTTGACGGTGAAACCCCAAATTGCCAG
CGCTAGTATTCTCATCAAATCACTTCCTCTTGGCTTCTCCATGCTGTATATT
35 CCGCGTGGACCAATCATGGATTACTCCAATCTAGATATTGTAACCTAAGGTC
CTTAAGGACCTTAAAGCTTTTGGCAAAAAACAAAGAGCTCTCTTTATCAAG
TGTGATCCTCTCATCTATTT

(SEQ ID NO: 112)

40 MYREITAVEHDRFVSESNQTNLLQSLNWPVKDNWGSQLLGFFDGETQIASA
SILIKSLPLGFSMLYIPRGPIMDYSNLDIVTKVLKDLKAFGKKQRALFIKCDPLI
Y

Sequence description:

5 A] Length: 326 bp - 108 aa (partial sequence)
 B] ATG start codon is preceded by an obvious
 Shine-Dalgarno sequence. No obvious leader
 peptide sequence.

10 ID-133

 Clone 2-c59

15 (SEQ ID NO: 113)

ATGGACAAGAAAAAAATCTTAGTAACGGGTATTGTGCCTAAAGAAGGTCT
AAGAAAGCTTATGGACCGATTTGATGTTACTTATTCAGAAGATCGCCCATT
TTCACGTGACTATGTGTTAGAGCATTTATCTGAATATGACGGATGGTTACT
CATGGGACAAAAAGGTGATAAAGAGATGATTGATGCAGGTGAAAACCTTAC
20 AAATTATTTCTTT

 (SEQ ID NO: 114)

MDKKKILVTGIVPKEGLRKLMDRFDVTYSEDRPFSRDYVLEHLSEYDGWLLM
25 GQKGDKEMIDAGENLQIIS

Sequence description:

30 A] Length: 215 bp - 71 aa (partial sequence)
 B] ATG start codon is preceded by an obvious
 Shine-Dalgarno sequence. No obvious leader
 peptide sequence.

35 ID-134

 Clone 2-c62

40 (SEQ ID NO: 115)

ATTTTCGAAAGATGACTACCAAAATATTAGTTTTTGGACAGGATCCAGAAGTT
GTTGATTATGCTGGTCTGTTTGAAAAACGCCGTCCAGTTTTAGAAAAAGCA
GTTAAAAATTTCTTGCAAGAAGAGAGAGCTACGAGAATGCTATCTGATTTC
TTGCAAGAAGAAAAATGGGTAACTGATTTTGCTGAATTTATGGCGATCAA

5 AGAACATTTTGGTAATAAGGCGCTTCAAGAATGGGATGACAAGGCTATTA
TACGCCGCGAAGAAGAAGCCTTAGCAGGATATCGTCAAAAGCTTAGTGAA
GTGATAAAATATCATGAAGTAACGCAATATTTCTTTTACAAACAATGGTTT
GAGTTAAAAGAATATGCTAATGATAAAGGGATTCAAATTATCGGTGATAT
GCCAATCTACGTTTCTGCCGATAGTGTAGAAGTTTGGACAATGCCTGAACT
GTTT

(SEQ ID NO: 116)
10 ISKDDYQNISFGQDPEVVVDYAGLFEKRRPVLEKAVKNFLQEERATRMLSDFLQ
EEKWVTDFAEFMAIKEHFGNKALQEWDDKAIIRREEEALAGYRQKLSEVIKY
HEVTQYFFYKQWFELKEYANDKGIQIIGDMPIYVSADSVEVWTMPELF

15 A) Length: 459 bp - 153 aa (partial sequence)
B) More sequencing is required to determine the
N- and C-termini
enzyme). - *Streptococcus pneumoniae* (63%)

20 ID-135

Identical to ID-108 described in WO 00/06736

25 Clone 2-c63

ID-136

30 Clone 2-c66

(SEQ ID NO: 117)
ATGGCAAAACAGAAAAATAACTGGCGCCGTGTTGGAGTTGGTGTCCTTAC
ACTTGCTTCAGTTGCGACTCTTGCTGCATGTGGAAGTAAATCAGCTTCCCA
GGATTCTAATGGAGCGATTAATTGGGCTATTCCAACAGAAATCAATACACT
35 AGATTTATCTAAAGTTACAGACACTTACTCAAATCTAGCTATTGGTAACTC
TAGTAGTAATTTTCCTTCGCTTAGATAAAGATGGAAAGACAAGACCAGACTT
GGCTACTAAAGTTGATGTTTCAAAGATGGCTTAAGTTATACAGCTACATT
ACGTAAAGGCTTGAAGTGGTCAGATGGCAGTAAAGTTACTGCAAAGGATT
TTGTTTATTCATGGCAACGTTTAGTTGATCCTAAAACAGCTTCACAATATG
40 CTTACCTTGCTGTTGAAGGGCATGTGCTTAATGCCGATAAAATCAACGAAG
GACAAGAGAAAGACTTGAATAAGCTAGGTGTTAAGGCAGAAGGCGATGA
CAAAGTTGTTATTACTTTATCTAGTCCGTCTCCGCAATTCATCTACTACCTT
GCATTCATAACTTCATGCCACAAAAACAAGAAGTTGTTGAAAAATATGG
AAAAGATTACGCAACTACTTCAAAAAATACAGTTTACTCAGGACCATATA

CTGTTGAAGGTTGGAATGGTTCGAATGGTACTTTCACGCTGAAGAAAAAC
AAAAATTATTGGGACGCTAAAAATGTAAAAACAAAAGAAGTTCGCATCCA
GACTGTTAAAAAACCAGATACCGCCGTTCAAATGTATAAACGTGGTGAGT
TAGATGCAGCTAATATCTCAAATACTTCTGCTATTTATCAAGCTAATAAAA
5 ATAATAAAGATGTCACAGATGTTCTAGAAGCGACCACTGCCTATATGGAA
TATAATACTACTGGTTCTGTGAAAGGGCTTGATAATGTTAAGATTCGTCGC
GCCTTAAACTTAGCAACTAACCGTAAAGGAGTTGTTCAAGCAGCCGTTGAT
ACAGGCTCAAAACCGGCAATTGCTTTTGCACCTACTGGTTTAGCCAAAACA
CCAGATGGAACTGATTTGGCAAAATATGTTGCCCCAGGTTATGAATATAAT
10 AAAACTGAAGCAGCAAAACTCTTTAGACTA

(SEQ ID NO: 118)
MAKQKNNWRRVGVGVLTLASVATLAACGSKSASQDSNGAINWAIPTEINTLD
LSKVTDITYSNLAIGNSSSNFLRLDKDGKTRPDLATKVDVSKDGLTYTATLRKG
15 LKWSDGSKLTAKDFVYSWQRLVDPKTASQYAYLAVEGHVLNADKINEGQEK
DLNKLGVKAEGDDKVVITLSSPSPQFIYYLAFTNFMPQKQEVVEKYGKDYAT
TSKNTVYSGPYTVEGWNGSNGTFTLKKKNKNYWDAKNVKTKEVRIQTVKKPD
TAVQMYKRGELDAANISNTSAIYQANKNNKDVTDVLEATTAYMEYNTTGSV
KGLDNVKIRRALNLATNRKGVVQAAVDTGSKPAIAFAPTGLAKTPDGTDLAK
20 YVAPGYEYNKTEAAKLFR

Sequence description:

25 A] Length: 1143 bp - 381 aa (partial sequence)
B] Shine-Dalgarno sequence precedes ATG codon.
Possesses a potential leader peptide sequence.

30 ID-137

Clone 2-c67

35 (SEQ ID NO: 119)
TTGAGAGTTTATGAAAATAAAGAAGAGTTGAAAAAAGAAATAAGTAAAAC
ATTTGAGAAATACATTATGGAATTTAATAA
TATTCCAGAGAATCTAAAAGATAAAAGAATTGATGAAGTTGATAGAACTC
CAGCAGAAAACCTTTCTTATCAGGTTGGCT
40 GGACCAACTTGGTTCTTAAATGGGAAGAAGATGAAAGAAAGGGACTTCAA
GTAAAAACACCATCGGATAAATTT

(SEQ ID NO: 120)

MRVYENKEELKKEISKTFEKYIMEFNNIPENLKDKRIDEVDRTPAENLSYQVG
WTNLVLKWEEDERKGLQVKTPSDKF

5

Sequence description

10 A] Length: 234 bp - 78 aa (partial sequence)
B] TTG start codon is preceded by a
potential Shine-Dalgarno sequence. No obvious
leader peptide sequence.

15 ID-138

Clone 2-c70

(SEQ ID NO: 121)

20 ATGTCAAAGTTTGATAGTCAGAAAATAATTACTCCGATTATGAAGTTTGTC.
AATATGCGAGGGATTATTGCACTCAAAGATGGCATGCTAGCAATTTTACCA
CTAACAGTTGTTGGGAGTCTCTTTTAAATATTAGGGCAGCTTCCATTT

(SEQ ID NO: 122)

25 MSKFDSQKIITPIMKFVNMRGILKDGMLAILPLTVVGSFLILGQLPF

Sequence description

30 A] Length: 150 bp - 50 aa (partial sequence)
B] ATG start codon is preceded by a potential
Shine-Dalgarno sequence. Possesses a potential
leader peptide sequence.

35

ID-139

Clone 2-c71

(SEQ ID NO: 123)

40 GAGACCACTTCATCAGTTAAACCAGCAGGAATTGACCGTATCAATCATACC
TCAACACCCCCGAAGAAAACCTACCCCCAACATTGCAACGACGCATAGCTT
CAAAGATCGTTGTGATACTTTAGAAAGAATTCACAATGAAGACATTGATGT
TTGTTCTGGATTCAATTTGTGGTATGGGAGAGAGCGATGAGGGGCTCATCAC

ATTAGCTTTCAGACTAAAAGAACTGAACCCCTATTCTATCCCTGTCAATTTT
T TACTTGCTGTTGAAGGAACACCTCTTGGAAAATATAACTATTTGACTCCC
ATTAAATGCTTAAAAATTATGGCCATGTTGCGTTTTTGTTTTTCCTTTCAAGG
AATTAAGATTAAGTGCTGGACGGGAGGTCCATTTTGAGAATTTTGAATCAT
5 TAGTCACCTTACTTGTTGACTCAACTTTTTTTGGGAAATTACCTAACAGAGG
GGGGTCGCAATCAACATACCGATATTGAATTCTTGGAAAAATTACAATA
AATCATACTAAAAAGGAATTAATTT

(SEQ ID NO: 124)
10 ETTSSVKPAGIDRINHTSTPPKKTTPNIAATTHSFKDRCDTLERIHNEIDVCSGFI
CGMGESDEGLITLAFRLKELNPYSIPVNFLLA VEGTPLGKYNLYLTPIKCLKIMA
MLRFVFPFKELRLSAGREVFHFNESLV TLLVDSTFLGNYLTEGGRNQHTDIEF
LEKLQLNHTKKELI

15
Sequence description:

20 A] Length: 535 bp - 178 aa (partial sequence)
B] N- and C-termini require verification

ID-140

25 Clone 2-c73

(SEQ ID NO: 125)
ATGCCGGTTTGGACTGCACAGTCTATTCCAAAGGCATTTTTAGAAAAGCAT
AATACTAAGGAAGGCACCTGGGCAAACTAACCATTCTAAGTGGTTCTTTA
30 GTATTTTACCAGTTATCTCCTGATGGAGAGGAAATCTCGCGGCATATTTTT
GATGCTAGTAGTGATATTCCTTTTTGTTGATCCACAAGTCTGGCATAAAGTT
TCGCCGAATAGTCCAGACTTAAGTTGCTATCTAACTTTTTTACTGCCAAAAA
GAAGATTACTTCCATAAAAAAATATGGTCTCACGCGCACACATTCTGAGGTT
ATCGCCAGTGCACCTCTCTTATCTGAGAAGAGTAATATATTAGACCTTGGG
35 TGTGGTCAAGGGCGAACTCACTTTATTTATCGCTGCTGGGACATCAAGTG
ACTTCTGTCGATTCAAACGGACAGAGCCTTG TAGCTTTAGAAAATATGGCA
TTAGAAGAAGAGCTTCCTTACAATATAAAAAGGTATGATATTAATACTACT
GCTATTGAAGGGCACTATGATTTTATTTTATCAACTGTGGTATTTATGTTTT
T

40
(SEQ ID NO: 126)
MPVWTAQSIPKAFLEKHNTKEGTWAKLTILSGSLVFYQLSPDGEEISRHIFDAS
SDIPFVDPQVWHKVSPNSPDLSCYLTFYCQKEDYFHKKYGLTRTHSEVIASAP

LLSEKSNILDLGCGQGRNSLYLSLLGHQVTSVDSNGQSLVALENMALEEELPY
NIKRYDINTTAIEGHYDFILSTVVFMF

5 Sequence description:

A] Length: 563 bp - 187 aa (partial sequence)
B] N- and C-termini require verification

10

ID-141

Clone 2c76

15

(SEQ ID NO: 127)

20

ATGACAAAGCAAATAATTGCCATTTGGGCTGAAGATGAAGACCATTTGAT
TGGAGTTAATGGCGGTTTACCATGGAGGCTTCCTAAAGAGTTACATCACTT
CAAAGAAACGACCATGGGGCAGGCTTTGCTTATGGGACGAAAGACCTTTG
ATGGAATGAACCGTCGTGTTTTACCTGGTAGAGAGACAATCATCTTAACAA
AAGATGAACAATTCCAAGCAGATGGAGTGACAGTCCTAAATAGTGTTGAA
CAAGTTATAAAATGGTTTCAGGAACATAATAAGACCTTATTTATTGTAGGT
GGTGCAAGTATTTATAAAGCATTCTGCCTTATTGTGAAGCAATCATAAAA
ACTAAAGTTCATGGAAAATTCAAAGGTGATACCTATTTTCCTGATGTTAAT
CTATCTGAGTTT

25

(SEQ ID NO: 128)

30

MTKQIIAIWAEDEDHLIGVNGGLPWRLPKELHHFKETTMGQALLMGRKTFDG
MNRRLVLPGRETIILTKDEQFQADGVTVLNSVEQVIKWFQEHNKTLFIVGGASI
YKAFLPYCEAIKTKVHGKFKGDTYFPDVNLSEF

Sequence description:

35

A] Length: 417 bp - 139 aa (partial sequence)
B] ATG start codon is preceded by a Shine-
Dalgarno sequence. No leader peptide sequence

40

ID-142

Clone 2-c78

(SEQ ID NO: 129)

TTGTGGCCAAACTGTGCCCCGCTTATTAATAGCACTTTGTTTCACCATTGAA
GATATCTTAACATCAGGTGCTCATAGCAACCCTATTTTAATGGGGGTTATA
CTTGGCGGGACAATTGTAGTAGTGGCGACAGCACCACCTTTCTTCTATGGCA
5 TTGACAGCTATGCTAGGATTAACCGGAATGCCTATGGCTATAGGAGCCTTG
TCTGTCTTTGGTTCGTCATTTATGAATGGTGTACTTTTCCATAAATTAAC
TTGGAAGTCGTAAAGATAATATAGCTTTTGCTGTTGAGCCTCTAACTCAAG
CTGACGTGACTTCAGCTAACCCTATTCCAATCTATGTCACTAATTTTGTTGG
TGGTGCAGCTTGTGGTATTTTAATTGCCTTGATGAAATTAGTTAATGATACT
10 CCTGGAACAGCGACACCAATTGCAGGATTTGCTGTCATGTTTGCCTATAAC
CCAATGATAAAAGTACTAATAACCGCTCTAGGTTGTATTATCCTATCTTTA
CTAGCAGGCTATTTTGGAGGCATTGTTTTT

(SEQ ID NO: 130)

15 MWPNCAPLINSTLFTIEDILTSGAHSNPILMGVILGGTIVVVATAPLSSMALTA
MLGLTGMPMAIGALSVFGSSFMNGVLFHKLKLGSRKDNIAFAVEPLTQADVT
SANPIPIYVTNFVGGAACGILIALMKLVNDTPGTATPIAGFVVMFAYNPMIKVL
ITALGCIILSLLAGYFGGIVF

20

Sequence description:

25 A) Length: 540 bp - 180 aa (partial sequence)
B) N- and C-termini have yet to be elucidated

ID-143

30 Clone 2-c80

(SEQ ID NO: 131)

ATGTTTTTAAGTATAATGGCAGGTGTCATAGCATTTGTCCTGACAGTTATT
GCCATTCCACGCTTCATTAAGTTTTACCAATTGAAGAAAATTGGCGGGCAA
35 CAAATGCATGAAGATGTCAAACAACATCTAGCCAAAGCAGGTACGCCGAC
AATGGGAGGAACGGTATTTT

(SEQ ID NO: 132)

40 MFLSIMAGVIAFVLTVIAIPRFIKFYQLKKIGGQQMHEDVKQHLAKAGTPTMG-
GTVF

Sequence description:

5 A] Length: 172 bp - 57 aa (partial sequence)
B] Shine Dalgarno sequence precedes 'ATG' start
codon. Possesses a potential leader peptide
sequence.

ID-144

10 Clone 3-83

(SEQ ID NO: 133)

15 ATGAAACCATATTTATCTTTTATTGGTAGAACGTTATTATACTTCGGTATTT
TATTGTTACTAATTTACTTTTTTGCATACCTTGGTCGCGGACAAGGCAGTTT
TATTTATAA

(SEQ ID NO: 134)

MKPYLSFIGRTLLEYFGILLLLIYFFAYLGRGQGSFIY

20

Sequence description:

25 A] Length: 113 bp - 37 aa (partial sequence)
B] Putative ATG start codon is preceded by a
typical Shine-Dalgarno sequence. Possesses a
potential leader peptide sequence.
This orf is not in frame with nuc

30

ID-145

Clone 3-86

35 (SEQ ID NO: 135)

40 ATGTCATATTTTAGAAATTACTGGTATCGTTTTGGAGCAATTTTATTTATTA
TTTAGCAGTAATATTGCTTGTTTTTAGACCTGACTGGTCAATGCTTCACTA
TCTATTGTATTTTACTTTATGGCACTTCTAGCGCATCAATTTGAAGAATAT
CAGTTTCCCGGTGGGGCATCACCTATCATTAAGTATGTTGTTTATGATGAA
GAAGAGCTGATGGATTGTTTTCCAGGCAATACTCAGTCTATTATGTTGGTT
AATACTATTGCTTGGTTGCTTTACATTGCTAGTATTGCTTTTCCTCAAGCTT
ATTGGCTTGGATTAGGAGTCATGTTCTTTAGTCTAACGCAGCTCTTGGGTC
ATGGTTTTTCAGATGAATATTAACTTAAACTTGGTATAATCCTGGTCTAG
CAACGACAGTATTTCTCCTAGTACCAATAGCTTGCGCATACATCTATCAAG

CTAGTGCAGAAGGAATGCTCACTTGGGGGAGATTGGCTAGGTGGTTTTATCA
TGTTGATTGTCTGTGTACTAACTAGCATTATTGCACCTGTACAGCTATTGAA
GGATAAGGAGACCAATTATATTATTAGTCCTTGGCAAATGGACCGTTTTCA
TAAGGTCGTTAATTTTGTAAGGATAAAAAAATAA

5

(SEQ ID NO: 136)

MSYFRNYWYRFGAILFIILAVILLVFRPDWSMLHYLLYFYFMALLAHQFEEYQ
FPGGASPIINYVVYDEEELMDCFPGNTQSIMLVNTIAWLLYIASIAFPQAYWLG
LGVMFFSLTQLLGHGFQMNIKLKTWYNPGLATTVFLLVPIACAYIYQASAEG
MLTWGDWLGGFIMLIVCVLTSIIAPVQLLKDKETNYIISPWQMDRFHKVVNFV
RIKK*

10

Sequence description:

15

A] Length: 651 bp - 219 aa (full length gene)
B] Putative ATG start codon is preceded by a
typical Shine-Dalgarno sequence. Possesses a
potential leader peptide sequence.

20

ID-146

25

Clone 3-c88

(SEQ ID NO: 137)

ATGCCACTTACAGCACTTGAAATTAAAGATAAAACATTTTCATCAAATTT
CGCGGTTATAGCGAAGAAGAAGTT

30

(SEQ ID NO: 138)

MPLTALEIKDKTFSSKFRGYSEEEV

35

Sequence description:

A] Length: 75 bp - 25 aa (partial sequence)
B] Putative ATG start codon is preceded by a
typical Shine-Dalgarno sequence. No leader
peptide

40

ID-147

Clone 3-90

(SEQ ID NO: 139)

5 ATGTCACCTTTTTCAAGAAAAAATTGCTTACAATTGCGCTAAAAAGGAAGCG
CTTTATAAAGAGAGTTTAGGACGCTACGCCTTGAGATCAATGCTAGCAGG
GGCTTATTTGACAATGAGTACTGCTGCCGGTATCGTCGCAGCTGATACTAT
TGGTAAAATTTCTCCTGCTCTATCAGGTTTTGTATTTGCTTTCATCTTTAGTT
TTGGACTTATTTATGTTTTAATATTTAATGGTGAATTGGCGACATCTAATAT
10 GCTTTATCTCACTGCAGGAGCCTATAATAAAAAATATCTCTTGGAAAAAAGC
CATAACAATTTTAATTTATTGTACTTTTTTCAACCTCGTTGGTGCTTGTATA
TTAGCTTGGTTGTTTAA

(SEQ ID NO: 140)

15 MSLFQEKIAYNCAKKEALYKESLGRYALRSMLAGAYLTMSTAAGIVAADTIG
KISPALSGFVFAFIFSGLIYVLIFNGELATSNMLYLTAGAYNKNISWKKAITILI
YCTFFNLVGACILAWLF

20 Sequence description

25 A] Length: 406 bp - 125 aa (partial sequence)
B] Putative ATG start codon is preceded by a
typical Shine-Dalgarno sequence. Possible
leader peptide

30 ID-148

Clone 3-92

(SEQ ID NO: 141)

35 AAGTTACAAGCGACTGAAGTTAAGAGCGTTCCGGTAGCACAACCAGCTTC
AACAACAAATGCAGTAGCTGCACATCCTGAAAATGCAGGGGCTCCAACCTC
ATGTTGCAGCTTATAAAGAAAAAGTAGCGTCAACTTATGGAGTTAATGAA
TTCAGTACATACCGTGCGGGAGATCCAGGTGATCATGGTAAAGGTTTAGC
AGTTGACTTTATTGTAGGTAAAAACCAAGCACTTGGTAATGAAGTTGCACA
GTA CTCTACACAAAATATGGCAGCAAATAACATTTTCAATATGTTATCTGGCA
40 AAAAAAGTTTTATTCAAATACAAATAGTATTTATGGACCTGCTAATACTTG
GAATGCAATGCCAGATCGTGGTGGCGTTACTGCCAACCCTATGACCACGT
TCACGTATCATTTAA

(SEQ ID NO: 142)

KLQATEVKSVPAQPASTTNAVAAHPENAGLQPHVAAAYKEKVASTYGVNEF
STYRAGDPGDHKGKGLAVDFIVGKNQALGNEVAQYSTQNMAANNISYVIWQQ
KFYSNTNSIYGPANTWNAMPDRGGVTANHVDHVHVSF

5

Sequence description

10

A] Length: 419 bp - 139 aa (partial sequence)
B] N- and C-termini have yet to be determined

ID-149

15

Clone 3-94

(SEQ ID NO: 143)

20 ATGATTCCAGTAGTTATTGAACAAACAAGTCGTGGTGAACGTTCTTATGAT
ATTTACTCACGTCTTTTAAAAGATCGTATTATTATGTTGACAGGCCAAGTT
GAGGATAATATGGCCAATAGTATCATTGCACAGTTATTGTTTCTCGATGCA
CAAGATAATACAAAGGATATTTACCTTTATGTCAATACACCAGGTGGTTCA
GTATCGGCTGGACTTGCTATTGTGGACACCATGAACTTCATTAAATCGGAC
GTACAGACGATTGTTATGGGGATGGCTGCTTCGATGGGAACCATTATTGCT
25 TCAAGTGGTGCTAAAGGAAAACGTTTTATGTTACCGAATGCAGAATATATG
ATCCACCAACCAATGGGCGGAACAGGCGGAGGTACACAGCAATCTGATAT
GGCTATCGCTGCTGAGCATCTTTTAAAACGCGTCATACTTTAGAAAAAAT
CTTAGCTGATAATTCTGGTCAATCTATTGAAAAAGTCCATGATGATGCAGA
GCGTGATCGTTGGATGAGTGCTCAAGAACACTTGATTATGGCTTTATTGAT
30 GCTATTATGGAAAATAATAATTTACAATAATAGATTTAAAAGAGTTGAGTT
TACCAACTCTTTTTTTTATTTGTTGGAATTATGTTATAATCTTAGTAATTACA
GATATGACGCAGAAAGGAAAAAATTATTGA

(SEQ ID NO: 144)

35 MIPVVIEQTSRGERSYDIYSRLLKDRIIMLTGQVEDNMANSIIAQLLFLDAQDN
TKDIYLYVNTPGGSVSAGLAIVDTMNFIKSDVQTIVMGMAASMGTIASSGAK
GKRFMLPNAEYMIHQPMGGTGGGTQQSDMAIAAEHLLKTRHTLEKILADNSG
QSIEKVHDDAERDRWMSAQEHLIMALLMLLWKIIYNNRFRVEFTNSFFICW
40 NYVIIIIVITDMTQKGKNY*

Sequence description

5 A] Length: 693 bp - 231 aa (full length gene)
 B] Putative ATG start codon is preceded by a
 typical Shine-Dalgarno sequence. No leader
 peptide. Significantly, it would appear to have a
 very hydrophobic C-terminus.

ID-150

10 Clone 2-c86

(SEQ ID NO: 145)
ATGAAACCAAAAaTTATTGGTGTACTTGGTCTAGGAATATTTGGACAAACA
CTCGCACAAGAACTAAGTAACTTTGAACAAGATGTTATTGCTATTGACAGC
15 AATCCTGAAAATGTACAAGCTGTCGCCGAAGT
TGTTACAAAAGCAGCTATCGGAGACATTACTGATTTAGCTTTCCTAAAACA
CATCGGGATCAGTGACTGTGATACTGTTATTATTGCTACAGGAAACAGTTT
AGAGAGCTCAGTATTGGCCGTAATGCACTGTAAAAAGTTAGGCGTCCCAC
AAGTTATTGCTAAAGCTCGAAACCTTGTATACGAAGAAGTACTTTATGAAA
20 TTGGTGCTGATTTGGTTATCTCTCCGGAGCGAGAATCTGGGCAAAAATGTTG
CTGCAAACCTCATGAGAAATAAAATTACAGATGTCTTCCAGATTGAATCTG
ATATTCTGTCATTGAATTT

(SEQ ID NO: 146)
25 MKPKIIGVLGLGIFGQTLAQELSNFEQDVIAIDSNPENVQAVAEVVTKAAIGDI
TDLAFLKHIGISDCDTVIIATGNSLE
SSVLAVMHCKKLGVPQVIKARNLVYEEVLYEIGADLVISPERESGQNVAAN
LMRNKITDVFQIESDISVIEF

30 Sequence description:

35 A] Length: 459 bp - 153 aa (partial sequence)
 B] Putative ATG start codon is preceded by a
 typical Shine-Dalgarno sequence. Possesses a
 potential leader peptide sequence.
 This orf is not in frame with nuc

40

ID-151

Clone 2-c88

(SEQ ID NO: 147)

GTGCGTTATAGTAAAGAGATTATTCAGTTAGCTATACCAGCTATGATTGAA
AATATCTTACAAATGCTCATGGGAGTAGTTGATAATTATCTAGTGGCTCAG
TTAGGTGTTGTAGCAGTATCAGGTGTTTCAGTTGCTAATAATATAATTACT
5 ATTTATCAAGCTATTTTTATAGCTTTAGGGGCGAGTATAGCAAGTCTATTG
GCCAAGTCGTTAGCAGGTAGTGAGAAGGATGATGCAATTTTCAGTATGTTCT
CAAGCCATTTTTCTAACATCACTGATAGGGGCAGTATTAGGAATTATCTCG
ATTGTTTTTGGACAAACTTTCTTT

10 (SEQ ID NO: 148)

MRYSKEIIQLAIPAMIENILQMLMGVVDNYLVAQLGVVAVSGVSVANNIITIY
QAIFIALGASIASLLAKSLAGSEKDDAISVCSQAIFLTSLIGAVLGIISIVFGQTF

15 Sequence description

A] Length: 330 bp - 110 aa (partial sequence)

20 B] Putative GTG start codon is preceded by a
typical Shine-Dalgarno sequence. May have a
leader peptide

ID-152

25

Clone 2-c92

(SEQ ID NO: 149)

30 TTGATTAACAAGTATTCGTGCTTTTTGAAGAGGATTCTCCATAATAATACT
CCTTTAATAGTTATCGTGAGAAGTATTTTAAAGAAAAACCGCCAAGGTAG
AGCGACATTTCTGCCTTTAACTACAATAAAACCAAGAGAATTAGCACAAAC
ATTATCTCTCAAAATTACAAAGTTCTCAAGGGTTTTTAGGAATAGCTAGTG
AATTGGTAACCTATGATCAACGCTTGTCAAACATTTTT

35 (SEQ ID NO: 150)

MINKYSCFLKRILHNNTPLIVIVRSILKKNRQGRATFLPLTTIKPRELAQHLYSK
LQSSQGFLGIASELVTYDQRLSNIF

Sequence description

40

A] Length: 240 bp - 80 aa (partial sequence)

B] No obvious Shine Dalgarno sequence precedes the Putative TTG start
codon

ID-153

5 Clone 2-c94

(SEQ ID NO: 151)

TTGTTGACTCACAAAAATATATTATTAACCATTATATTTGGATTATTTATGA
TTATATTATCAGCATGTGGTATGTCTAATAAGGAAATGGCTGGTATTGATA
10 ATTGGGAACATTATCAAAAAGGAAAAGAAAATTACTATTGGATTTGATAAT
ACTTTTGTTCCTATGGGATTTGAAAGTCGTTCTGGTGACTATACCGGCTTTG
ATATTGATTTAGCTAATGCTGTTTTTAAAGAATACGGTATTTTCAGTGAAAT
GGCAGCCTATTAACCTGGGATATGAAAGAACTGAACTTAATAATGGTAAT
ATAGACCTTATTTGGAATGGTTATTCAAAAACGGCAGAACGTGCTAAAAA
15 AGTCGCTTTTACAAACCCATATATGAATAATCATCAAGTAATTGTTACTAA
AACTTCATCACATATTAATAGTATTAAGGATATGAAGGGGAAAAAACTAG
GAGCCCAGTCGGGTTCATCTGGTTTTTGATGCTTTTAAACGCTAAACCTGATA
TTTTAAAAAAGTTTGTAAAAGGAAAAGAAGCAGTTCAATACGATACTTTC
ACTCAGGCTTTGATTGATTTAAAAAATAACCGTATTGATGGTCTTTTGATT
20 GATGAAGTTTATGCTAACTATTATTTAAAGCAAGAAGGAA

(SEQ ID NO: 152)

MLTHKNILLTIIFGLFMIILSACGMSNKEMAGIDNWEHYQKEKKITIGFDNTFV
PMGFESRSGDYTGFDIDLANAVFKEYGISVKWQPINWDMKETELNNGNIDLI
25 WNGYSKTAERAKKVAFTNPYMNNHQVIVTKTSSHINSIKDMKGKKLGAQSG
SSGFDAFNAKPDILKKFVKGKEAVQYDTFTQALIDLKNNRIDGLLIDEVYANY
YLKQEG

30 Sequence description

35 A] Length: 649 bp - 216 aa (partial sequence)
B] TTG start codon is preceded by a possible
typical Shine-Dalgarno sequence. Has a
leader peptide

40 ID-154

Clone 2-c100

(SEQ ID NO: 153)

ATGAAAATTTGGAAAAAATAACCTTAATGTTTTCTGCAATTATTTTAACA
ACAGTAATTGCATTGGGAGTCTATGTTGCCTCAGCTTATAATTTTTCGACTA
ATGAATTGTCTAAGACTTTT

5

(SEQ ID NO: 154)

MKIWKKITLMFSAILTTVIALGVYVASAYNFSTNELSKTF

10 Sequence description

15 A] Length: 123 bp - 41 aa (partial sequence)
B] ATG start codon is preceded by a potential
typical Shine-Dalgarno sequence. Has a
typical leader peptide

ID-155

20

~

Clone 2-c1

(SEQ ID NO: 155)

25 ATGAAAAACAAAGACTATTACTGCTTTTTGGAGGCTTATTAATAATGATA
ATGATGACAGCATGTAAGGATTCAAAAATCCCAGAAAACCGCACGAAAAA
GGAATACCAGGCAGAACAGAATTTTAAGTCATACTTTAAATATATATCAG
ATAAAAATAACTATTTAGATAATATAAAAGTTTATTACTTTTCTATAAGTA
TTTCTAAAGATGTACAAGATAAAGTCAGTGAAACAACAACCTTGTTTCATATA
GACTAGAAAAGCAAAAGAATCAAGAGTTCATTGGTAATTTTGAACATGAA
30 GTTAGTGAATCTAGTCAATATTCAACCGAAGTTAAAAATCAAATACAGTAT
CCAATCCAGTATAAAGATAATTCAATTCGTTTTACTGAAAAAACACCGTCA
GAACGTTATGATGAGTTTGTTTTTAGTTTCAATTTGATTCTTCATTATTAAAAA
AATATAAAATATATGATTACTTACTAAAACATCCCGAAACTGAATTAAAA
GGTGTTTCCTATAAGATTCCTATAAATTCTGAAATTGTAGCCCCTTTTATAA
35 ATCAATTAAATATAAAAAAATCCTAAAAAATCATCTATTTCGGTTACAAAAA
CGGAAAGTAAAGAATATTATTATACAATCAGTATTGATACTGATTCTGAGA
TATATTCTATATTCGAAGGTATTCAT

(SEQ ID NO: 156)

40 MKKQRLLLLFGGLLIMIMMTACKDSKIPENRTKKEYQAEQNFKSYFKYISDKN
NYLDNIKVYYFISISISKDVQDKVSETTTCSYRLEKQKNQEFIGNFEHEVSESSQ
YSTEVKNQIQYPIQYKDNSIRFTEKTPSERYDEFVFSSFDSSLLKKYKIYDYLLK
HPETELKGVSYKIPINSEIVAPFINQLNIKNPKKSSISVTKTESKEYYYTISIDTDS
EIYSIFEGIH

Sequence description

5

A] Length: 687 bp - 229 aa (partial sequence)
B] ATG start codon is preceded by a potential
typical Shine-Dalgarno sequence. Has a
typical leader peptide. C-terminus has yet to be
verified

10

ID-156

15

Clone 2-c5

(SEQ ID NO: 157)

ATGACATTTGACACCATTTGATCAATTAGCGGTTAATACAGTCCGCACGCTT
TCTATTGATGCTATCCAAGCAGCAAATTCTGGGCACCCAGGTCTTCCTATG
GGAGCTGCGCCTATGGCTTATGTGCTTTGGAATAAATTCTTAAATGTAAAC
CCAAAAACAAGTCGCAATTGGACAAACCGTGACCGTTTTGTACTTTCAGCT
GGGCATGGTTCAGCTCTTCTTTATAGCCTACTTCATTTAGCTGGCTATGATT
TATCAATTGATGATTT

20

25

(SEQ ID NO: 158)

MTFDTIDQLAVNTVRTLSIDAIQAANS GHPGLPMGAAPMAYVLWNKFLNVNP
KTSRNWTNRDRFVLSAGHGSALLYSLHLAGYDLSIDD

30

Sequence description

35

A] Length: 272 bp - 90 aa (partial sequence)
B] ATG start codon is preceded by a potential
typical Shine-Dalgarno sequence. No obvious
leader peptide

40

ID-157

Clone 2-c8

(SEQ ID NO: 159)
ATGAGAACACTATTTAGAAATGATATTTGCTATTCCAAAGTTTATCTTTAGA
TTGATTTGGAATATCATTGTTGGGAATATTCAAGACAGTTCTTGTTATTGCG
ATTATTTTATTTGGCTTGTATTACTATGCGAATCACAGTCAATCAGAATTG
5 CTAATCAACTTAGTGACATTATTCAGACAGGAAAAACATTTT

(SEQ ID NO: 160)
MRTLFRMIFAIPKFIFRLIWNIIWGIFKTVLVIAIILFGLYYYANHSQSEFANQLS
DIIQTGKTF
10

Sequence description

15 A] Length: 197 bp - 65 aa (partial sequence)
B] ATG start codon is preceded by a potential
typical Shine-Dalgarno sequence. Possesses a
leader peptide

20 ID-158

Clone 2-c9

25 (SEQ ID NO: 161)
ATGTCAAAAAAAAAATAATATTAGGAATTTTATCTCTTTTATCTGTCGTTACTT
TGGTGGCGTGTGGTTCATCAGACAAACAGCTACAAGATAAAGTTGAGAAA
AAAGGGAAGTTAGTTTTAGCGGTGAGTCCAGATTATGCTCCCTTTGAGTTT

30 (SEQ ID NO: 162)
MSKKIILGILSLLSVVTLVACGSSDKQLQDKVEKKGKLVLA VSPDYAPFEF

35 Sequence description

40 A] Length: 153 bp - 51 aa (partial sequence)
B] ATG start codon is preceded by a potential
typical Shine-Dalgarno sequence. Possesses a
leader peptide (not in frame with nuc)

ID-159

Clone 2-c10

(SEQ ID NO: 163)

5 ATGAAAAATCAAAGACTATTACTGCTTTTTGGAGGCTTATTAATAATGATA
ATGATGACAGCATGTAAGGATTCAAAAATCCCAGAAAACCGCACGAAAAA
GGAATACCAGGCAGAACAGAATTTTAAGTCATACTTT

(SEQ ID NO: 164)

10 MKNQRLLLLFGGLLIMIMMTACKDSKIPENRTKKEYQAEQNFKSYF

Sequence description

15

A] Length: 139 bp - 46 aa (partial sequence)
B] ATG start codon is preceded by a potential
typical Shine-Dalgarno sequence. Possesses a
leader peptide

20

ID-160

Clone 2-c11

25

(SEQ ID NO: 165)

ATGATTGGAAAATTATATTATAGCTATAGAAAGTCACGCTTATTAAGAAGT
ATTTTATGGCTTATTTTAATTGTTGGTGTATATATGTTAGGACAACGTGTTT
TATTATCCACTGTTTCCTTTATCACATCAAGAGATAAACTAGCAGTAGATC
30 AACATTTACTCAATAACTTTTCAGCAGTAAGTGGTGGGAGTTTTAATAAAT
TAAATGTTTTTCACACTGGGGTTGAGTCCATGGATGTCAAGTATGATTATTT
GGAGATTCGTTTCCTTATTTTCGTGGGCAAAAAAATGCAACGAAGCGAAAA
GCAGAAGTAGCTCAATATACTTTAATGCTTACTATCTCAGTTATACAAGCA
TATGGTGTTTCAGGAAATCAATTTATAAAAAGCTCTTTATTAGGTTCTTATA
35 GTGATATTGTTTTT

(SEQ ID NO: 166)

MIGKLYYSYRKSRLRLSILWLILIVGVYMLGQRVLLSTVPLSHQEIKLAVDQHL
LNNFSAVSGGSFNKLNVFTLGLSPWMSSMIWRFVSLFSWAKNATKRKA-EVA
40 QYTLMLTISVIQAYGVSGNQFIKSSLLGSYS-DIVF

Sequence description

5 A] Length: 423 bp - 141 aa (partial sequence)
B] ATG start codon is preceded by a potential
typical Shine-Dalgarno sequence. Possesses a
leader peptide

ID-161

10 Clone 2-c13

(SEQ ID NO: 167)
ATGAAAGGTCTATTGGATTTTTTTAGTTAATATTGCCAGAACGCCAGCTATT
TTAGTCGCCTTGATAGCCATTATCGGTTTAGTACTGCAGAAAAAAGGTGTT
15 CCTGATATTGTAAAAGGTGGAATAAAAACATTTGTTGGCTTCTTAGTGGTT
TCTGAAGGTGCAGGGATAGTCCAAAATTCCTTGAATCCATTTGGAAAAATG
TTTGAACATGCTTTTCATTTGGTGGGGGTAGTTCCTAATAATGAAGCCATT
GTAGCAGTAGCTCTTACGAAGTATGGCTCAGCAACTGCTTTGATTATGTTA
GCGGGAATGATTTTTTAATATTTTAATTGCTCGTTTTACAAAA

20 (SEQ ID NO: 168)
MKGLLDFLVNIARTPAILVALIAIIGLVLQKKGVDPDIVKGGIKTFVGFLVVSEG
AGIVQNSLNPFGKMFHAFHLVGVVPNNEAIVAVALTKEYGSATALIMLAGMI
FNILIARFTK

25

Sequence description

30 A] Length: 348 bp - 116 aa (partial sequence)
B] ATG start codon is preceded by a potential
Shine-Dalgarno sequence. Possible leader
peptide

35

ID-162

Clone 2-c21

40 (SEQ ID NO: 169)
TTGGTTGGTAAGCCCCAATTACTATTTTTAGATGAACCTACTTCCGGAATG
GATACTTCCACACGTCAACGATTTTGGGAAGCTGGTTGCGACACTAAAAAA
AGAAGGTGACACAATTGTCTATTCTAGTCATTATATCGAAGAGGTAGAAC
ATACAGCTGATAGGATTTTAGTACTTCATAAAGGAAAGTTATTACGCGATA

5 CAACCCCCTTTGCCATGAAGCAAGAAAAAACCGAAAAGTTATTCACCGTT
CCGCTTAGTTATCAAAAATTATTACCTACCTATTTGATTACAGAGTGTGAA
GCCAAGAGTGATAGTATAACGTTTGTTACTGGGGAGGCTGAAACTGTATG
GAAAATACTGGCAGATAATGGTTGTCCTATTGAAGCTATTGAGATGACCA
ATAGAACTTTGTAAATCGTATTTTTTGAGACTACTAAGGAGGTAAAACATG
AGAATCTTTA

(SEQ ID NO: 170)
10 MVGKPQLLFLDEPTSGMDTSTRQRFWKL VATLKKEGDTIVYSSHYIEEVEHTA
DRILVLHKGKLLRDTPFAMKQEKTEKLFTVPLSYQKLLPTYLITECEAKSDSI
TFVTGEAETVWKILADNGCPIEAIEMTNRTLNRIFETTKEVKHENL

15 Sequence description

A] Length: 462 bp - 155 aa (partial sequence)
B] B] Putative TTG start codon is not preceded by
20 an obvious Shine-Dalgarno sequence. No obvious
leader peptide. N- and C- termini require further
examination.

25 ID-163

Clone 2-c25

(SEQ ID NO: 171)
30 TTGAAAAAATCCAAGAGAAGCCGTAAGGCAGTGACAACAAGTGGTGAGA
AGACTTTACTTGAGGATTTGGCAAAAATGAATTCCTAGACGAAGTCATTA
ATGTTATGGTTTTATATACCTTGAATAAGACAAAATCTGCTAACTTAAATA
AGGCCTATATCATGAAAGTTGCTAATGATTTTGCCTTTCAGAATGTTATGA
CGGCCGAAGATGCTGTGCTTAAAATTCGTGATTTTTTCAGATCAAAAAGTAA
GGACTAAAACAGAAACGAAGAAGAAACAATCGAATGTTTCCTGAATGGAGT
35 AATCCTGATTATAAAGATGAGGTTAGCCCAGAAAAAAGAAATTGAATTAGA
ACAGTTT

(SEQ ID NO: 172)
40 MKKSKRSRKAVTTSGEKTLLEDLAKMNFLDEVINVMVLYTLNKTKSANLNK
AYIMKVANDFAFQNVMTAEDAVLKIRDFSDQKVRTKTETKKKQSNVPEWSN
PDYKDEV SPEKEIELEQF

Sequence description

A] Length:360 bp - 120 aa (partial sequence)
B] N- and C- termini require verification.

5

ID-164

Clone 2-c28

10

(SEQ ID NO: 173)

ATGACGAATCATATTACTAAACTGATAGAAAATAGCGGAAAAAAATTGAC
AGAAATTAGCGAAGCTACAGATATAGCCTATCCTACACTTTCTGGATACAA
TCAAGGAATCCGCAAACCTAAAAAAGATAATGCTGAAAAAATTGGCAAAAT
ACTTTAATGTTTCCGTCGCTTACATTATGGGACTTGATAGCAACCCACATG
CTCCATCAAATCTT

15

(SEQ ID NO: 174)

MTNHITKLIENSGKKLTEISEATDIA YPTLSGYNQGIRKPKKDNAEKLAKYFNV
SVAYIMGLDSNPHAPSNL

20

Sequence description

25

A] Length:218 bp - 72 aa (partial sequence)
B] ATG start codon is preceded by an
obvious Shine Dalgarno sequence. No obvious
leader peptide.

30

ID-165

Clone 2-c29

35

(SEQ ID NO: 175)

TTGATGAAAAGGAATAAACATTTACCGTTAACAGAACTACCTATTATATT
TTATTAGCTTTGTTTGAGGAAGCGCATGGCTATTATGAAAAAAGTT
GAAGAAATGAGTGGCGGTGATGTTAGAATAGCCGCAGGGACAATGTACGG
TGCCATTGAAAATTTACTTAAACAAAAAATGGATAAAGTCTATCTCAAGTGA
CGATAGAAGAAGAAAAGTTTATATTATTACTGAGACAGGAAAAGAAATAG
TAGAACTTGAAACGAATCGATTAAGAAAGTTACTTAATACTGCTAATCAGT
TGGGTTTTGGAGGAGATGGTTATGATAAAGTTT

40

(SEQ ID NO: 176)

MMKR NKHLPLTET TYYILLALFEEAHGYAIMKKVEEMSGGDVRIAAGTMYG
AIENLLKQKWIKSISSDDRRRKVYIITETGKEIVELETNRLRKLLNTANQLGFG
GDGYDKV

5

Sequence description

10

A] Length:337 bp - 112 aa (partial sequence)

B] TTG start codon is preceded by an
obvious Shine Dalgarno sequence. Actual start
codon may ATG that comes immediately after the
TTG. Potential leader peptide.

15

ID-166

Clone 2-c35

20

(SEQ ID NO: 177)

CCCATTACTGGTGAGTTAATAGCTGAGAAATTAGGAGTACCAAGAGCAGC
ACTAAGGTCTGATTTGCGGGTTTTAAGTATGCTAGGTATCATAGATGCAAA
ACCTAAGGTTGGTTATTTTTATTTAGGACAGTATCATGCTTCAATAGGGAC
25 AAGTCATTTTGAAAAGATGACAGTTTCAGAAATTATGGGGATCCTTCTGAC
AGTTCATCAAAAAGATTCAGTTTATGATGTTATTGTACATATTTTTATGGA
AGATGCTGGTTGTGCTTTTATCTTGGATGATGATGATTTTCTCTGTGGAGTC
GTGTCACGTAAAGATTTACTAAAAACCAGTATTGGCGGAGGAGATCTTTCT
AAAATGCCAATAGGAATGGTGATGACACGTATGCCACACGTGACAACTGT
30 TTTAGAAAATGAAAGTCTTTTTGCGGCAGCTGATAAATTAGTGAGCAGAA
AAGTGGATAGTCTCCCTGTCGTTTCGTCATGATAAGCAATATCCCGAAAAAT
TTA

30

(SEQ ID NO: 178)

PITGELIAEKLGVPRALRSDLRVLSMLGIIDAKPKVG YFYLGQYHASIGTSHF
35 EKMTVSEIMGILLTVHQKDSVYDVIVHIFMEDAGCAFILDDDDFLCGVVSRKD
LLKTSIGGGDLSKMPIGMVMTRMPHVTTVLENESLFAAADKL VSRKVDSL PV
VRHDKQYPEKF

35

40

Sequence description

A] Length:511 bp - 170 aa (partial sequence)

B] N- and C-termini to be determined

ID-167

5

Clone 2-44

(SEQ ID NO: 179)

10 TTGGAAGTCATCATGCAATTTATTTATAGTATTATTGGTATTTTATTGGTAT
TAGGAATTGTGTATGCAATTTCTTTCAATCGTAAGAGTGTTTCTCTAAGTTT
AATTGGAAAAGCTCTTATCGTTCAATTCATTATTGCGCTAATCTTAGTACGT
ATCCCACTAGGCCAACAAGTTGTTAGTGTTGTTTCAACTGGAGTTACTAAA
GTAATCAACTGTGGTCAAGCTGGTTT

15

(SEQ ID NO: 180)

MEVIMQFIYSHIGILLVLGIVYAISFNRKSVSLSLIGKALIVQFIILVRIPLGQQ
VVSVVSTGVTKVINCGQAG

20

Sequence description

25

A] Length:233 bp - 77 aa (partial sequence)
B] TTG start codon is preceded by a
possible Shine Dalgarno sequence. Actual start
codon may occur further downstream. Potential
leader peptide.

30

ID-168

Clone 2-46

(SEQ ID NO: 181)

35 CAACCTAATAAAGCTTTAGAAAGTGATGAGATTGATATTAATGCTTTCCAG
CATTATAATTACTTAACCAATTGGAATAAAGCAAATAAGACCAATCTTGTT
TCCGTTGCTGAGACATACTTTACTTCCTTTAGATTATACTCTGGTACTAAGA
ACGGTAAAGGTAAATACCAAACAGTTTCTGAAATTCCAAATAAAGCAACT
ATTACTATCCCAAACGATGCAGTTAACGAAAGTCGCTCTCTCTACTTGTTA
40 CAATCAGCAGGCTTGCTAAAATTGAAAGTATCAGGTGATACATTAGCAAC
AATGTCAGATGTTGTTTCCAATCCTAAATCTTTAGATTT

(SEQ ID NO: 182)

QPNKALESDEIDINAFQHYNLTNWNKANKTNLVSVAETYFTSFRLYSGTKN
GKGKYQTVSEIPNKA TITIPND AVNESRSLYLLQSAGLLKLKVSGDTLATMSD
VVSNP KSLD

5

Sequence description

10

A] Length:344 bp - 114 aa (partial sequence)
B] N- and C- termini require verification

ID-169

15

Clone 2-47

(SEQ ID NO: 183)

20

ATGAAATGTATAATAAATAATATAAATAAAAATGATAATTGAGAT
TTATCATAGAAGGAAAAC TATTTTGAAATTAAATAAAAATCATATTATCTAC
TGCAGCTCTTACTGCTCTCTTTT TAGGATATAATAGCGTTACTGCGGATACA
TATAATAACTATCAGCCACATAGATCAAATAATATGGATT TAACTGAGGA
ATATAACTATAATAACCAGATAGAACTTCAGGAGCGTATAAAAAAACCTAA
ATATACCTTTT

25

(SEQ ID NO: 184)

MKCIINNINKIKMII EIYHRRKTILKLNKIILSTAALTALFLGYNSVTADTYNNY
QPHRSNNMDLTEEYN YNNQIELQERIKNLNIPF

30

Sequence description

35

A] Length:264 bp - 88 aa (partial sequence)
B] There is a Shine-Dalgarno sequence upstream
of this sequence. Potential leader peptide
sequence

40

ID-170

Clone RS-58b

(SEQ ID NO: 185)

TTGGGTGATTATTATGGTAAGAAATATTTTGGTGAGGCAGCTAAAAAAGA
CGTCGAACATATGGCTAAGAAAATCATTAATGTCTATAAAACACGGTTAA
AAAACAACACTTGGTTATCAGAAAATACAAAAGCAATGGCCATTAAGAAA
5 CTTGATAACATGAGATTAATGATTGGCTATCCAGAAGATTATCCTGATCTT
TATCGTCAGTACCAATTTGATAGTAAAGCAAGCTTCTTTGAAAACAATGAT
AACTACAGAAAATTATCGAACAAGAAAACATTTGAAGAATTTAACCAGTC
TAATCAACGTGAACATTGGCAAATGAGTGCCAATGCTGTAAATGCTTATAA
TGATCCTAATAACCAATTCCATAGTCTTTCCAGCAGCGATTTTTCAATCACCA
10 CTGTACGATAAAACTAAAACAGTTAGTCAAAATTATGGAGCTATCGGAGC
AATTATTGGTCATGAAATTTACACTCATTTGATATTAATGGTATGAAATA
TGACGAGAAAGGGAATCTTCACGATTGGTGGACTAAAGAAGATTTAAATC
ATTATAAGAAATCAACACAAGCTATGATTGACCAATGGGATGGCCTTAAA
GCAGATGGCGGTAAAGTTGATGGTAAATTAACTTTAGCAGAAAATATTGC
15 AGATAATGGTGGTGTATGGCATCTCTAGAAGCTCTTAAGACTGAAAAAAT
CCAAACTATAAAGAATTTTTTTGAATCATGGGCAAGTATTTGGCGTCAAAAA
GCAACCAAAGAACAAAGTAAGTCCTCAATTCAGTCAGATGTTTCATGCACC
ATATGAATTGAGAGCTAACATCCCAGTACGTAATTTCCAAGAATTTTATGA
TGCCTTTGGTGTAAAAAAGGCGATTCAATGTATCTAAAACCAGAAAAAC
20 GTTTGACACTTTGGTAA

(SEQ ID NO: 186)

MGDYYGKKYFGEAAKKDVEHMAKKIINVYKTRLKNNTWLSENTKAMAIKK
LDNMRLMIGYPDYPDL YRQYQFDSKASFFENNDNYRKLSNKKTFEEFNQSNQ
25 REHWQMSANAVNAYNDPNTNSIVFPAAIFQSPLYDKTKTVSQNYGAIGAIIGH
EISHSFDINGMKYDEKGNLHDWWTKEDLNHYKKSTQAMIDQWDGLKADGG
KVDGKLTLAENIADNGGVMASLEALKTEKIQTIKNFLNHGQVFGVKKQPKNK
VSPQFSQMFMHHMN*

30 Sequence description:

A) Length: 819 bp - 272 aa (full length gene)
(107 bp of additional DNA sequence (> onwards) is
also included. While not in-frame with the
35 described orf, it also shares strong homology
with the neutral peptidases.

B) This gene sequence was not identified using the LEEP system. It was
identified downstream of the ID-89 gene which was identified by LEEP,
during cloning and sequence analysis of the full-length ID-89 gene sequence.
40 ID-89 and ID-170 together show homology over their combined entire length
with the neutral endopeptidases from Lactococcus and Lactobacillus. Possesses
TTG (possible ATG start codon located 13 bp further downstream) start
codon with no obvious signal peptide. Shine Dalgarno sequence not
immediately obvious. Possibly located further downstream

ID-171

5 Clone 2-18/22b (Mod2)

(SEQ ID NO: 187)

ATGACCATGATTACGCCAAGCTTCATTAAGGTATCTCTAGATGAAACAAAT
CGTATGATGCGTATGATATCAGATTTATTAAGTTTATCGCGCATTGATAAT
10 GAAGTAACGCATTTAGATGTTGAAATGACGAATTTTACAGCTTTCATGACC
TCAATTTTGAATCGATTTGATCAGATTAGAAATCAAAAAACAGTCACAGG
AAAAGTTTATGAAATTGTCAGAGATTATCCTCTTAAGTCAATTTGGGTGGA
AATTGATACAGATAAGATGACTCAAGTGATTGATAACATTTTAAATAATGC
AGTCAAGTATTCACCAGATGGTGGTAAGATTACAGTTAATCTACGCACAAC
15 TAAAACGCAGATGATTTTATCAATATCAGACCAAGGCTTAGGTATTCCCAA
AAAAGATTTACCTCTCATTTTTTGATCGTTTTTATCGTGTTGATAAGGCGAGA
AGTCGTCAACAGGGTGGGACTGGACTTGGTTTGTCAATTGCAAAAGAAAT
TGTTAAGCAGCATAAGGGATTTATTTGGGCTAAGAGTGAGTATGGTAAAG
GGTCTACTTTTACAATCGTCTTGCCTTATGATAAAGATGCTGTAACCTTATGA
20 AGAATGGGAGGACGTTGAAGATTAA

(SEQ ID NO: 188)

MTMITPSFIKVSLDETNRMMRMISDLLSLSRIDNEVTHLDVEMTNFTAFMTSIL
NRFDQIRNQKTVTGKVYEIVRDYPLKSIWVEIDTDKMTQVIDNILNNAVKYSP
25 DGGKITVNLRTTKTQMILSISDQGLGIPKKDLPLIFDRFYRVDKARSRQQGGTG
LGLSIAKEIVKQHKGFIWAKSEYGKGSTFTIVLPYDKDAVTYEEWEDVED*

Sequence description:

30 A] Length: 613 bp - 212 aa (full-length gene possibly)
B] Possible Shine Dalgarno sequence present
upstream of a ATG start codon. May not have yet
determined the N- portion of this gene. No
35 obvious signal peptide.

ID-172

40 Clone 2-54balternate (107b)

(SEQ ID NO: 189)

TTGAAAAAAATTATTACTTCTATTCTATTACTTAGTTGCATTTTTTTTATGC
CAACCATCTCTGCTGAATCTTTTAATGCTTCCGCTAAACATGCCTTAGCAGT

5 TGATTTAGATTTCAGGAAAAATCTTGTATGAAAAAGATGCTAACAAACCCG
CTGCTATTGCTTCCTTGACTAAAATAATGACCGTTTATATGGTCTATAAAG
AAATTGATAACGGTAACCTCAAGTGGAATACCAAAGTAAATATATCTGAC
TACCCTTATCAACTAACACGCGAATCTGATGCTAGTAATGTTTCCTTTAGAA
AAAAGGCGCTATACTGTAAACAACCTCGTGGACGCTGCCATGATTTCTAGT
GCTAACAGTGCAGCCATTGCTTTAGCTGAACATATTTTCAGGAACTGAAAGT
AAATTTGTTGATAAAATGACTGCTCAATTGGAAAAGTGGGGAATTCATGAT
AGCCACCTAGTCAATGCTTCTGGCTTAAATAAATAGTATGTTAGGCAATCAC
ATTTATCCAAAATCGTCACAAAACGACGAAAATAAAAATGAGTGCACGTGA
10 TATTGCTATTGCTGCCTACCATTGTTGGTCAACGAATATCCTTCCATTCTTAAG
ATTACTAGTAAGTCCGTTGCTAAATTTGATAAAGATATTATGCATTCTTAT
AACTACATGCTACCAGATATGCCTGTCTTTAGACCAGGTATTACAGGTTTG
AAAAGTGGGACAACGGAATTAGCTGGCCAATCTTTTATTGCTACATCTACT
GAAAGTGGAAATGAGACTACTCACTGTTATTATGCATGCTGATAAGGCCGAT
15 AAAGACAAATATGCTCGCTTTACAGCAACTAAGTCTCTCTTGAAGTATATC
ACAAACACCTACGAACCTAACCTTGTATTAGCTAAAGGAGCTGCATATAA
AGGTAAAGAAGCAAGTGTGAGAGACGGAAAAGAACAATCGGTCATCGCT
GTTGCTAAAAACGATTTGAAAGTAGTACAGAAGAAAAAATATCACTAAACA
AAATCAGTTAAAAATTAACCTTTAAAAAAGAGCTTACTGCTCCTATTACAAA
20 AAAAGAGAACCTAGGGAAAGCTTATTACGTTGACCTTAATAAGGTTGGAA
AAGGCTATCTCATAAAGGAACCTAGCGTTCATTTAGTGGCAAAAGATAGT
ATTGAGCGCAGTTTCTTCCTCAAAGTGTGGTGGAAATCATTTTGTGCGCTAC
GTTAACGAAAAACTTTAA

25 (SEQ ID NO: 190)
MKKIITSILLSCIFFMPTISAESFNASAKHALAVDLD SGKILY EKDANKPAAIA
SLTKIMTVYMVYKEIDNGNLKWNTKVNISDYPYQLTRESNASNP LEKRRYT
VKQLVDAAMISSANSAAIALAEHISGTESKFVDKMTAQLEKWGIHDSHLVNA
SGLNNSMLGNHIYPKSSQNDENKMSARDIAIAAYHLVNEYPSILKITSKSVAKF
30 DKDIMHSYNYMLPDMPVFRPGITGLKTGTTEL AGQSFIATSTESGMRLLTVM
HADKADKDKYARFTATNSLLNYITNTYEPNLVLAKGAAYKGKEASVRDGKE
QSVIAVAKNDLKVVQKKNITKQNQLKINFKKELTAPITKKENLGKAYYVDLN
KVGKGYLIKEPSVHLVAKDSIERSFFLKVWWNHFVRYVNEKL*

35 Sequence description:

A] Length: 1236 bp - 412 aa (full-length gene sequence possibly)

B] A possible Shine-Dalgarno sequence precedes the putative 'TTG' start
codon. (needs further cloning and sequencing to verify N-terminus)

40

ID-173

Clone 3-60b

(SEQ ID NO: 191)

5 ATGACGCTTCGAGAATTAACAATAGAAGAATTTAAAGAACATTCAGGAAA
TTATGATTCACAATCATTTTTACAAACACCTGAGATGGCTAAACTTTTAGA
AAAACGCGGCTATGATGTTAGGTATTTGGGATATCAAGTAGAAAATAAAC
TAGAGATAATCAGTTTATCTTATATTATGCCAGTCACTGGTGGTTTTCAAAT
GAAAATTGATTCAGGACCAGTTCATTCAAATTCTAAGTATCTAAAACAATT
TTATAAAGCATTGCAAGGCTATGCCAAATCCAACGGTGTTCTAGAATTAAT
AGTTGAGCCTTTTGATGATTACCAATTATTCAGTTCGGGAGTTCCTAGT
10 AATCAGGGAAATGATAATCTGATTGAAGATTTTACCAGTTCAGGTTATCAC
CATGATGGTTTAACAACCTGGTTTTACTGGTAAATATTTATCTTGGCACTATG
TTAAAAATTTAGAAGGTGTCACCTTCTGAAACGTTACTATCTTCATTCTCTAA
GACAGGACGAGCTTTGGTTAAGAAAGCAATGTCTTTTGGAATCAAGGTTC
GCGTTCTTAAACGTGATGAGCTACATTTATTTAAAGAGATAACAACCTTCTA
15 CGTCAAATAGACGTGATTATATGGATAAGTCCTTAGATTATTATCAAGATT
TTTACGATAGCTTTGAAGGCAAGGCTGAATTTGTGATTGCCACTTTAAATT
TTAGAGAATACGACCATAACTTGCAAATAAAAGCTGAAGCATTGGAAAAT
AAGCTT

20

(SEQ ID NO: 192)

MTLRELTIEEFKEHSGNYDSQSFLQTPEMAKLLEKRGYDVRYLGYQVENKLEI
ISLSYIMPVTGGFQMKIDSGPVHSNSKYLKQFYKALQGYAKSNGVLELIVEPF
DDYQLFTSSGVPSNQGNLIEDFTSSGYHHDGLTTGFTGKYLSWHYVKNLE
25 GVTSETLLSSFSTGRALVKKAMSFGIKVRVLKRDELHLFKEITTSTSNRRDY
MDKSLDYYQDFYDSFEGKAEFVIATLNFREYDHNLQIKAEALENKL

Sequence description

- 30 A) Length: 771 bp - 257 aa (partial gene sequence)
B) This gene sequence was not identified using the LEEP system. It was
identified immediately downstream of the ID-65 gene which was identified by
LEEP, during cloning and sequence analysis of the full-length ID-65 gene
sequence. Sequence Characteristics:
35 No obvious leader peptide sequence
Orf is preceded by a potential Shine-
Dalgarno sequence.

40 ID-174

Clone 2-17b (ID-80b)

(SEQ ID NO: 193)

TTGTCATTAAGTTTGGTTGCAGTGTTAAATCTTATCCCTCCTAAAATCATGG
GATCAGTTATTGATGCTATTACAACCTGGAAAATTAACAAGACCACAATTAC
TATGGAATTTATTAGGTTTGGTTTTTGTGAGCTTTAGCTATGTATGGGCTGCG
TTATATTTGGCGTATGTATATTTTAGGGACTTCTTACAAATTAGGCCAAGTT
5 GTCAGATACCGTTTATTTGAACATTTTACAAAAATGTCTCCTTCTTTTTATC
AGAAATATCGTACAGGTGATTTAATGGCGCACGCGACCAACGACATCAAT
TCTCTAACACGTCTTGCAGGAGGAGGAGTTATGTCAGCAGTGGATGCCTCT
ATCACAGCATTAGTAACGCTTATCACCATGTTCTTTACTATTTTCGTGGCAA
ATGACATTAATTGCGGTTATCCCTTTGCCCTTAATGGCCTTAGCACTAGTA
10 AATTGGGGCGAAAAACCCATGAAACCTTCAAAGAATCTCAGGCAGCCCTT
TTCAGAATTAATAATAAAGTG

(SEQ ID NO: 194)

MSLSLVAVLNLIPPKIMGSVIDAITTGKLTRPQLLWNLLGLVLSALAMYGLRYI
15 WRMYILGTSYKLGQVVRYLFEHFTKMSPSFYQKYRTGDLMAHATNDINSLT
RLAGGGVMSAVDASITALVTLITMFFTISWQMTLIAVIPLPLMALALVNWGEK
PMKPSKNLRQPFSELNNKV

20 Sequence description

A) Length: 534 bp - 178 aa (partial gene
sequence)

B) This gene sequence was not identified using the LEEP system. It was
identified downstream of the ID-80 gene which was identified by LEEP,
25 during cloning and sequence analysis of the full-length ID-80 gene sequence.
Sequence Characteristics:
No obvious leader peptide sequence
Orf is preceded by a potential Shine-
Dalgarno sequence.

30

ID-175

Clone 2-11Ab (ID-103b)

35

(SEQ ID NO: 195)

ATGCATATTGAGACTGTTATTGATTTCAAAGAATTAGGAAAAAGATATCGT
TTTAAAAATCCTACAAAAGAATTAATAGCTGATACTTTAGAACAAGTCTTA
GAAGTGATAAAAGAAGTTGATTATTATCAATCTCAAAATTATTATGTTGTT
40 GGTTATTTATCTTATGAAGCATCTGCTGCTTTTGATTACATTTTAAAGTTT
CTCAACAGAAGTTGGCTGGAGAACATCTAGCTTATTTTACAGTACATAAAG
ATTGTGAGAACGAAGCTTTTCCTTTAAGTTATGAAAATGTTAGATTAGCAG
ATAATTGGACTGCTAATGTTTCTGAGCAAGAATATCAAGAGGCAATTGCTA

ATATTAAAGGACAAATTAGACAAGGAAATACTTATCAAGTAAATTATACA
CTAGAGCTTAGCCAACAATTATGCTCGGATCC

(SEQ ID NO: 196)

5 MHIETVIDFKELGKRYRFKNPTKELIADTLEQVLEVIKEVDYYQSQNYVVG
LSYEASAAFD SHFKVSQQKLAGEHLAYFTVHKDCENEAFPLSYENVRLADNW
TANVSEQEYQEALANIKGQIRQGNTYQVNYTLELSQQLCSD

10 Sequence description:

A] Length: 440 bp - 146 aa (partial gene sequence)

15 B] This gene sequence was not identified using the LEEP system. It was
identified downstream of the ID-103 gene which was identified by LEEP,
during cloning and sequence analysis of the full-length ID-103 gene sequence.
Shine Dalgarno sequence present upstream of
ATG start codon, No apparent leader peptide sequence

20 ID-176

Clone 2-18/22b(b) (ID-104b)

(SEQ ID NO: 197)

25 GTGAATAATATGTTTTATCTCAAAATAGCCTGGCATAATTTAAAACATTCT
ATAGACCAGTACATAACCATTCCTTAGCCAGTTTATTACTTTATTCATTGA
CTTGTTCTACGCTACTAATCTTAATGAGTGCTGTTGGAAGAGATATGGGGA
CAGCGGCAACGGTTCTTTTTCTTGGAGTGATTGTTTTGTCAATCTTTGCGGT
AGTCATGGAACATTATAGCTACAATATCTTGATGAAACAGCGTAGTAGTG
30 AATTTGGACTGTATAACATTTTGGGGATGAATAAACGTCAAGTTGCGCGTG
TAGCTAGTCTAGAGCTGTTTATTATTTATATATTTCTTATTTCTATAGGAAG
TCTGTTTAGTGCTTTTTTTTGCTAAATTTATTTATTTAATTTTTGTCAACATTA
TTAACTATCATGCACTAAATCTTAGTTTAAGTTTATGGCCATTTATTATTG
TATCGTTATATTTACAGGTATTTTTCTGACTTTAGAAGTTCCAGTTATTCGA
35 CATGTTCAATTTATCATCCCCATTAAGTCTTTTTAGAAAGAAACAACAGGGA
GAAAAAGAACCACAAAAGGTAATCTTATACTTGCAATTTTAGCGTTAGTAGCT
ATCGCCATCGCTTATACAATGGCTCTTACTTCAGGTAAAGCACCTGCATTA
GCTGTTATCTATCGTTTCTTCTTTGCAGTACTTTTAGTAATTGCTGGTACTT
ATCTTTTTTATATTAGTTTTATGACATGGTACTTAAAAAGGTTGCGTCAAAA
40 CAAGCATTATTATTATAAATCTGAGCATTTTGTATCAACTTCGCAAATGAT
TTTTCGAATGAAGCAAAAATGCAGTAGGGTTAGCAAGTATCACTTTATTAGC
TGTTATGGCTCTAGTTACTATTGCTACAACAGTCTCACTCTATTCAAATACA
CAAAATGTTGTTACCGGACTATTTCCAAAATCAGTAAGTTTATCAATAGAT
AATTCAAAAGGTGACGCGAAAAATATATTTGAAGAAAAGATTTTGAAGAA

ACTAGGTAAGTCATCTAAGGAAGCTATCACTTATAATCAGACAATGATTTC
GATGCCAGTTAGTCAATCAAGTGACTTAATATCACATCTA

(SEQ ID NO: 198)

5 MNNMFYLKIAWHNLKHSIDQYIPFLLASLLLYSLTCSTLLILMSAVGRDMGTA
ATVFLFLGVIVLSIFAVVMEHYSYNILMKQRSSEFGLYNILGMNKRQVARVASL
ELFIIYIFLISIGSLFSAFFAKFIYLIFVNIINYHALNLSLSLWPFIIICIVIFTGIFLTLE
VPVIRHVHLSSPLSLFRKKQQGEKEPKGNLILAILALVAIAIAYTMALTSGKAP
10 ALAVIYRFFFAVLLVIAGTYLFYISFMTWYLRRLRQNKHYYYKSEHFVSTSQM
IFRMKQNAVGLASITLLAVMALVTIATTVSLSYNTQNVVTGLFPKSVSLSIDNS
KGDANKIFEEKILKKLGKSSKEAITYNQTMISMPPVSQSSDLISHL

Sequence description:

15 A] Length: 1119 bp - 373 aa (partial gene sequence)
B] This gene sequence was not identified using the LEEP system. It was
identified upstream of the ID-104 gene which was identified by LEEP, during
cloning and sequence analysis of the full-length ID-104 gene sequence.
20 Possible Shine Dalgarno sequence present
upstream of a GTG start codon. Possesses a potential
leader peptide sequence

25 ID-177

Clone 2-5b (ID-112b)

(SEQ ID NO: 199)

30 ATGGTTGAGCCAATTATTTCAATACAAGGACTTCATAAAAGTTTTGGGAAA
AATGAGGTTTTAAAAGGCATTGACTTGGATATTCATCAAGGAGAAGTGGT
GGTTATTATTGGCCCTTCTGGCTCTGGTAAGTCAACATTTTTAAGAACAAT
GAATCTCTTGGAAGTACCAACAAAGGGAACAGTGACTTTTGAAGGGATTG
ATATAACAGACAAAAAGAATGATATTTTTAAAATGCGCGAAAAAATGGGC
35 ATGGTTTTTCAACAGTTCAATCTATTTCCCAATATGACTGTACTAGAAAAT
ATTACTTTATCACCTATTAAGACAAAGGGACTTTCTAAGCTTGATGCTCAG
ACAAAAGCATACGAGCTACTTGAAAAAGTTGGACTCAAAGAGAAGGCTAA
TGCTTATCCAGCAAGCTTATCTGGAGGACAACAACGATTGCTATTGC
AAGAGGTCTTGCAATGAATCCTGATGTCCTTCTTTTGGATGAACCTACTTCA
40 GCTCTTGATCCTGAAATGGTAGGTGAAGTCTTGACTGTTATGCAAGATTTA
GCTAAATCTGGTATGACGATGGTTATTGTCACTCATGAAATGGGTTTTGCA
CGTGAAGTAGCGGATCGTGTCAATTATTATGGATGCAGGGATTATTGTTGAG
CAAGGGACCCCTAAGAAAGTATTTGAGCAGACAAAAGAAATCCGCACAAG
AGACTTCTTAAGTAAAGTATTATAA

(SEQ ID NO: 200)

MVEPIISIQGLHKSFGKNEVLKGIDLDIHQGEVVVIIGPSGSGKSTFLRTMNLLE
VPTKGTVTFEGIDITDKKNDIFKMREKMGMVFQQFNLFPMNTVLENITLSPIKT
5 KGLSKLDAQTKAYELLEKVGLKEKANAYPASLSGGQQORIAIARGLAMNPDV
LLFDEPTSALDPEMVGEVLTVMQDLAKSGMTMVIVTHEMGFAREVADRVIF
MDAGIIVEQGTPKKVFEQTKEIRTRDFLSKVL*

10 Sequence description:

A] Length: 735 bp - 244 aa (full length gene)

15 B] This gene sequence was not identified using the LEEP system. It was
identified downstream of the ID-112 gene which was identified by LEEP,
during cloning and sequence analysis of the full-length ID-112 gene sequence.
Shine-Dalgarno sequence precedes the 'ATG'
start codon. No obvious leader peptide

20 ID-178

Clone 2-5c (ID-112c)

(SEQ ID NO: 201)

25 ATGTCTCA^sTATCAAGAGTGGTTAGAAAACGACTCACTCGGTAAAGATATT
AAGTCAGATTTAGAAGCTATTAAAGGAGATGAATCTGAAATTCAGGATCG
TTTTTACAAAACATTAGAATTTGGAACGGCGGGATTGAGAGGTAAACTTG
GAGCAGGAACCAATCGTATGAATACTTATATGGTGGGGAAAGCAGCACAA
GCATTAGCTAATCGATTATTGATCATGGCCCTGAAGCTATTGCACGTGGAA
30 TTGCAGTTAGTTATGATGTCCCGTTATCAATCTAAGGAATTTGCAGAATTA
ACTTGGTCCATTATGGCAGCAAATGGTATTAAAGCCTTATATTTA

(SEQ ID NO: 202)

35 MSHMNYKEIYQEWLEND^sLGKDIKSDLEAIKGDESEIQDRFYKTLEFGTAGLR
GKLGAGTNRMNTY^sMVGKAAQALANRLLIMALKLLHVELQLVMMSRYQSKE
FAELTWSIMAANGIKALYL

Sequence description:

40

A] Length: 366 bp - 122 aa (partial gene sequence)

B] This gene sequence was not identified using the LEEP system. It was
identified downstream of the ID-112 gene which was identified by LEEP,

during cloning and sequence analysis of the full-length ID-112 gene sequence. Shine-Dalgarno sequence preceded the 'ATG' start codon. No obvious potential leader peptide sequence.

5

ID-179

Clone 2-5d (ID-112d)

10

(SEQ ID NO: 203)

15

20

25

30

35

40

ATGCAACCTGTAAAAGTCGATGAACCTTCTGTTGAAGAAACCATTACTATT
TTGAAAGGTATCCAAAAAATAACGAAGATTATCATCACGTAAAATATAA
TAATGATGCCATAGAAGCAGCTGCAGTACTATCTAATCGTTATATCCAAGA
CCGCTTTTTACCTGATAAAGCAATAGACTTATTAGATGAAGCTGGTTCTAA
AATGAACCTAACACTAAATTTTGTGATCCAAAAGAAATTGATCAACGTCT
CATTGAAGCAGAAAATTTAAAAGCGCAAGCGACTCGTGAAGAAGATTACG
AACGTGCAGCTTACTTCCGTGACCAGATTGCAAAAATATAAAGAAATGCAG
CAACAAAAGGTCGACGATCAAGATACACCTATTATTACCGAAAAAACAAT
TGAGCACATCATTGAAGAAAAAACGAATATCCCTGTTGGTGATTTAAAAG
AAAAAGAACAATCTCAATTAATTAATCTCGCAGATGACTTGAAACAGCAT
GTGATCGGCCAGGATGACGCTGTCATTAAGATTGCAAAAAGCTATTCGTCGT
AATCGAGTTGGTCTTGGTAGCCCAAACCGTCCTATTGGTTCCTTTTTATTG
TAGGACCAACCGGTGTTGGTAAACTGAACTTTCTAAACAACCTAGCAATTG
AGCTCTTTGGTTCAGCTGATAGTATGATTCGTTTTGATATGTCAGAGTACAT
GGAAAAGCATGCTGTTGCTAAATTAGTCGGAGCGCCTCCAGGATACGTGG
GATACGAGGAAGCTGGACAACCTAACTGAAAAGGTTTCGTGCAAATCCTTAC
TCGCTCATCCTTCTAGATGAAATTGAAAAAGCTCATCCCGATGTCATGCAT
ATGTTCTTGCAGGTCCTTGATGACGGTCGATTAACAGATGGACAAGGAAG
AACTGTTAGTTTTAAAGATACCATTATCATCATGACCTCAAATGCTGGTTC
TGGTAAACTGAAGCAAGTGTTGGCTTTGGTGCCTCACGAGAAGGTAGGA
CGAATTCGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCAT
GCAAGC

(SEQ ID NO: 204)

MQPVKVDEPSVEETITILKGIQKKYEDYHHVKYNNDAlEAAAVLSNRYIQDRF
LPDKAIDLLDEAGSKMNLTLNFVDPKEIDQRLIEAENLKAQATREEDYERAAY
FRDQIAKYKEMQQQKVDDQDTPIITEKTIEHIIIEKTNIPVGDLKEKEQSQLNL
ADDLKQHVIGQDDAVIKIAKAIRNRVGLGSPNRPISFLFVGPTGVGKTELSK
QLAIELFGSADSMIRFDMSEYMEKHAVAKLVGAPPGYVGYEEAGQLTEKVRR
NPYSLILLDEIEKAHPDVMHMFLLQVLDDGRLTDGQGRTVSFKDTIIMTSNAGS
GKTEASVGFASREGRTNSSSVPGDPLESTCRHAS

Sequence description:

A] Length: 1070 bp ÷ 356 aa (Partial gene sequence)

B] This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-112 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-112 gene sequence. Shine-Dalgarno sequence preceded the 'ATG' start codon. No obvious potential leader peptide sequence.

5

10

ID-180

Clone 2-7b (ID-113b)

15

(SEQ ID NO: 205)

20

ATGAGAGGGAAGGTTATTTACGGCACAACCCTTATAGGTCTTTTTCTATTC
TTATTTTTCTATTTTTGGATTCTAAGCATCACATCGAGAGAATACATCATC
ATCGTATAAAGCAGGTAGATGCGAAGAGTGATTTAACAGGATTTAAAACC
CATTTGCCCATTATCAGCATTGATACAAAGCAACAAGTTATTCCTCTTGTT
ACAAAAGAAGGCGGAAAATATGTCAAAGCTAGGGATAATATTAATGTTGA
TATCGAATTACGGGATTCTCCAAGTAGATCACATCATTTATCAGAAAAGCC
GAGAATTAGGACAAAAGGGTTAATATCATATAGAGGAAATTCCTCTCGTT
ACTTTGATAAGAAGTCATTGAAAGTTAAGTTTGTTACTAATAAGTTAAAGG
AAAAGAAGCATCGATTAGCAGGAATGCCTAAAGAATCGGAGTGGGTATTG
CATGGTCCCTTTCTAGACAGAACATTATTAAGAAATTATCTGAGTTATAAT
ATTGCTGGTGAGATTATGCCTATGCCCAAACGTTTCGCTACTGTGAGTTAT
TTGTCAATGGTGAGTATCAGGGAG

25

(SEQ ID NO: 206)

30

MRGKVIYGTTLIGLFLFLFFYFWIPKHHIERIHHHRIKQVDAKSDLTGFKTHLP
SIDTKQQVIPLVTKEGGKYVKARDNINVDIELRDSPSRSHHLSEKPRIRTKGLIS
YRGNSSRYFDKKSLKVKFVTNKLKEKKHRLAGMPKESEWVLHGPFLDRTLRL
NYLSYNIAGEIMPMPQTFATVSYLSMVSIRE

35

Sequence description:

A] Length: 582 bp - 194 aa (Partial gene sequence)

B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-113 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-113 gene sequence. ATG start codon is preceded by a Shine-Dalgarno sequence-Possesses a potential leader peptide sequence. C-terminus to be determined.

40

ID-181

5 Clone 2-17b (ID-117b)

(SEQ ID NO: 207)

CTTCACATTTTATTGATCACTATCTGACAAATGTTAATCAAACAGCAGTTCT
TATTTTAGTGGGATATTATTCAATGTATGTCTTGCAGACCTTAATTCAATAT
10 TTTGGGAATCTCTTTTTTTCGCGCGTGTTTCTTATAGTATTGTTAGAGATATTC
GTAGAGATGCTTTTGCTAATATGGAAAGGCTAGGCATGTCTTATTTTGATA
GGACACCGGCAGGATCTATTGTGTCACGTATTACTAATGATACTGAAGCAA
TATCTGATATGTTTTTCGGGTATTTTATCAAGTTTTATCTCGGCGATATTTAT
TTTTACAGTTACTCTGTACACTATGTTGATGCTAGACATTAACTAACAGG
15 ACTCGTCGCTCTTTTGTTACCTGTTATCTTTATATTAGTGAATGTCTATCGG
AAAAAATCAGTCACTGTCATTGCTAAAACGAGAAGTTTACTTAGTGATATC
AACAGTAAATTATCAGAAAGTATTGAAGGAATTC

(SEQ ID NO: 208)

20 SHFIDHYLTNVNQTA VLILVGYYS MYVLQTLIQYFGNLFFARVSY SIVRDIRRD
AFANMERLGMSYFDRTPAGSIVSRITNDTEAISDMFSGILSSFISAIFIFTVTLYT
MLMLDIKLTGLVALLPVIFILVNVYRKKSVTVIAKTRSLSDINSKLSEIEGI

25 Sequence description:

A] Length: 498 bp - 165 aa (Partial gene sequence)

30 B] This gene sequence was not identified using the LEEP system. It was
identified downstream of the ID-117 gene which was identified by LEEP,
during cloning and sequence analysis of the full-length ID-117 gene sequence.
N- and C-termini have yet to be determined

ID-182

35

Clone 3-8b (ID-120b)

(SEQ ID NO: 209)

40 ATGTACCATATTGAATTAAAAAAGGAAGCTTTACTACCAAGAGAAACGCCT-
AGTTGATTTAGGCGCAGATAGATTGAGTAATCAGGAGTTATTAGCCATTCT
CTTACGTACAGGTATTAAAGAAAAACCTGTTCTTGAAATTTCAACGCAAAT
TTAGAAAACATAAGCAGTTTAGCAGATTTTGGTCAATTATCCTTACAGGA
GTTGCAATCCATTAAAGGAATCGGTCAGGTAAATCCGTCGAAATAAAAG
CTATGCTAGAACTAGCAAAACGGATTCACAAAGCTGAATATGATCGTAAA

5 GAGCAAATTTTAAGTAGTGAACAATTAGCGAGGAAAATGATGCTCGAATT
AGGGGATAAAAAACAAGAACATTTAGTAGCTATTTATATGGATAACACAAA
ATCGTATTATCGAACAGAGAACTATTTTTATTGGTACTGTACGTCGTTTCAG
TAGCAGAGCCAAGAGAAATTCTACATTATGCTTGTA AAAACATGGCAACT
TCTTTGATTATTATACATAATCATCCCTCAGGTTCTCCAAATCCCAGTGAAA
GTGATTTAAGTTTCACTAAAAAAATAAAACGATCATGTGATCATCTGGGAA
TTGTCTGCCTAGATCACATCATCGTTGGAAAAAATAAATATTATAGTTTTTC
GAGAAGAAGCAGATATTTTATAA

10 (SEQ ID NO: 210)
MYHIELKKEALLPRERLVDLGADRLSNQELLAILLRTGIKEKPVLEISTQILENI
SSLADFGQLSLQELQSIKGIGQVKSVEIKAMLELAKRIHKA EYDRKEQILSSEQ
LARKMMLELGDKKQEHLVAIYMDTQNRHIEQRTIFIGTVRRSVAEPREILHYAC
KNMATSLIIH NHPSGSPNPSESDLSFTKKIKRSCDHLGIVCLDHIIVGKNKYYSF
15 REEADIL*

Sequence description:

20 A) Length: 681 bp - 227 aa (full-length gene)
B) This gene sequence was not identified using the LEEP system. It was
identified downstream of the ID-120 gene which was identified by LEEP,
during cloning and sequence analysis of the full-length ID-120 gene sequence.
ATG start codon is preceded by an typical
25 Shine-Dalgarno sequence. No obvious leader
peptide sequence

ID-183

30 Clone 3-11b (ID-121b)

(SEQ ID NO: 211)
TGGTTAAAAGTAGTAGCTTGTATTCCATCTATTTTAATTGCTTTACCAT
35 TTGATAATTGGTTTGAAGCTCATTTTAATTTCATGATTCCGATTGCAATAGC
CCTAATCTTTTATGGTTTTGTCTTCATATGGGTTGAAAAACGTAATGCACAC
CTCAAACCACAGGTAACCGAATTGGCAAGTATGTCTTACAAGACAGCTTTC
TTGATTGGATGTTTCCAGGTTCTCAGTATTGTTCCGGGAACCAGTCGTTCTG
GAGCTACTATTTTAGGAGCAATTATTATTGGAAGTAGTCGTTCCGGTCGETG
40 CTGACTTTACTTTCTTCCTTGCCATCCCAACTATGTTTGGTTATAGTGGACT
TAAGGCGGTAAATATTTTTTAGATGGTAACGTCTTGAGTTTAGACCAATC
TTTAATACTTTTAGTAGCAAGTCTGACAGCTTTCGTAGTTAGTTTATATGTT
ATTCGTTTCTTGACAGACTATGTCAAACGACACGATTTCACCATCTTTGGT

AAGTATCGTATAGTCTTAGGAAGTTTACTCATCCTCTACTGGTTAGTTGTTT
ATTATTCTAA

(SEQ ID NO: 212)

5 WLKVVIACIPSILIALPFDNWFEAHFNFMIPALALIFYGFVFIWVEKRNAHLKP
QVTELASMSYKTAFLIGCFQVLSIVPGTSRSGATILGAIIGTSRSVAADFTFFLA
IPTMFGYSGLKAVKYFLDGNVLSLDQSLILLVASLTAFVVSLEYVIRFLTDYVKR
HDFTIFGKYRIVLGSLLILYWLVVHLF*

10

Sequence description:

A] Length: 579 bp - 193 aa (partial sequence)

15 B] This gene sequence was not identified using the LEEP system. It was
identified downstream of the ID-68 gene which was identified by LEEP,
during cloning and sequence analysis of the full-length ID-68 gene sequence
described in WO 00/06736. N-terminus has yet to be determined.

20

ID-184

Clone 3-11c (ID-121c)

(SEQ ID NO: 213)

25 ATGGAAATGAAACAAATCAGTGAAACAACACTGAAAATTACAATTAGTAT
GGAAGATTTAGAAGATCGTGGTATGGAGCTGAAAGATTTCTAATCCCTCA
GGAGAAGACTGAGGAATTTTTCTATTCTGTCATGGATGAATTAGACTTGCC
AGAAAACTTTAAAAAATAGTGGTATGTTAAGTTTTTCGAGTAACACCTAAAA
AAGATCGCATTGATGTTTTTTGTTACAAAGTCTGAATTAAGTAAAGATTTAA
30 ATTTAGAAGAATTAGCAGATTTGGGTGACATTTCAAAAATGTCTCCAGAAG
ACTTTTTTAAAACCTTGGAACAATCGATGTTGGAAAAAGGGGATACGGAT
GCCCATGCCAAATTAGCAGAAATTGAAAATATGATGGATAAAGCAACTCA
AGAAGTAGTTGAGGAAAATGTTTCTGAAGAACAACCTGAAAAGGAAGTAG
AAACGATTGGATATGTTCACTATGTCTTTGATTTTGATAATATTGAAGCTGT
35 AGTTCGATTTTCACAAACGATTGATTTTCCAATAGAAGCTT

(SEQ ID NO: 214)

MEMKQISETTLKITISMEDLEDRGMELKDFLIPQEKTEEFFYSVMDELDLPENF
KNSGMLSFRVTPKKDRIDVFVTKSELKDLNLEELADLGDISKMSPEDFFKTLE
40 QSMLEKGD TDAHAKLAEIENMMDKATQE VVEENVSEEQPEKEVETIGYVHY
VFDFDNIEAVVRFSQTIDFPIEA

Sequence description:

5 A] Length: 547 bp - 182 aa (Partial sequence)
B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-68 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-68 gene sequence. ATG start codon is preceded by an typical Shine-Dalgarno sequence. No obvious potential leader peptide sequence

10

ID-185

15 Clone 3-16b (ID-122b)

(SEQ ID NO: 215)

GGAAACCAACGGCCAGTACAATCGTCAAGGGTAGATTATCCTAAACGTAG
TCGTGCCAAGATTGTAGAAGTTTATTTTAGACAAGCTTCTACTACTGATTA
TTCTGGTGTTTACAAAGGTTACTATATTGACTTTGAAGCCAAAGAAACCCG
20 GCAGAAAACCTGCTATGCCTATGAAAAATTTTCATGCTCACCAAATAGAGC
ACATGGCAAATGTATTACAGCAAAAAGGGATTTGCTTTGTCTTGCTTCATT
TTTCCACACTTAAGGAAACCTATCTACTCCCTGCTAATGAGTTAATTTTCATT
TTATCAGATTGATAAAGGCAATAAATCAATGCCTATTGATTATATCAGAAA
AAATGGATTTTTCGTAAAGGAGAGTGCCTTTCCTCAAGTCCCTTACTTAGA
25 TATTATTGAAGAAAAATTATTAGGCGGTGATTACAATTAA

(SEQ ID NO: 216)

GNQRPVQSSRVDYPKRSRAKIVEVYFRQASTTDYSGVYKGYIIDFEAKETRQ
KTAMPMKNFHAHQIEHMANVLQQKGICFVLLHFSTLKETYLLPANELISFYQI
30 DKGNKSMPIDYIRKNGFFVKESAFPQVPYLDIIEEKLLGGDYN*

Sequence description:

35 A] Length: 447 bp - 149 aa (partial sequence)
B] This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-122 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-122 gene sequence. N-terminus has yet to be determined

40

ID-186

Clone 3-17b (ID-123b)

(SEQ ID NO: 217)

5 GGATCCTAAAAACGCTAAGGTTTATCAAAAAAATGCTGATCAATTTAGTG
ACAAGGCAATGGCTATTGCAGAGAAGTATAAGCCAAAATTTAAAGCTGCA
AAGTCTAAATACTTTGTGACTTCACATACAGCATTCTCATACTTAGCTAAG
CGATACGGATTGACTCAGTTAGGTATTGCAGGTGTCTCAACCGAGCAAGA
ACCTAGTGCTAAAAAATTAGCCGAAATTCAGGAGTTTGTGAAAACATATA
AGGTAAAGACTATTTTTGTTGAAGAAGGAGTCTCACCTAAATTAGCTCAAG
CAGTAGCTTCAGCTACTCGAGTTAAAATTGCAAGTTTAAGTCCTTTAGAAG
10 CAGTTCCCAAAAACAATAAAGATTACTTAGAAAATTTGGAAACTAATCTTA
AGGTACTTGTCAAATCGTTAAATCAATAG

(SEQ ID NO: 218)

15 DPKNAKVYQKNADQFSDKAMAIAEKYKPKFKAASKYFVTSHTAFSYLAKR
YGLTQLGIAGVSTEQEPSAKKLAEIQEFVKTYKVKTIFVEEGVSPKLAQAVAS
ATRVKIASLSPLEAVPKNNKDYLENLETNLKVLVKS LNQ*

Sequence description:

20 A] Length: 433 bp - 144 aa (partial sequence)
B] This gene sequence was not identified using the LEEP system. It was
identified downstream of the ID-123 gene which was identified by LEEP,
during cloning and sequence analysis of the full-length ID-123 gene sequence.
25 N-terminus has yet to be determined

ID-187

30 Clone 3-46/47 (ID-130b)

(SEQ ID NO: 219)

ATGAAAAAAGTCATCGATTTAAAAAACTACAAAAAGCATAACGCCTCAGA
AACTGTTTTAAATAATATTAATTTGGAGGTGTTTAAAGGAGAAATAATTGG
35 ATTAATAGGACCCTCTGGAGCAGGGAAATCTACCTTGATTA AAACTATGCT
TGGCATGGAAAAAGCAGATAAGGGAACAGCTCTTGTTCTTGATACTCAA
TGCCAGATCGTAATATTTTAAATCAAATTGGCTATATGGCTCAATCTGATG
CCTTACACGAGTCTTTAACTGGCTTAGAAAATTTATTATTCTTTGGAAAAA
TGAAAGGTATTCAAAAAAAGTGAATTA AACAGCAGATAAATCATATTTCT
40 AAAGTAGTAGATCTAGAAAACCAACTTGATAAATTTGTCTCAGGTTACTCA
GAAGGTATGAAAAGACGGCTTTCTCTAGCCATCGCCCTACTTGGAACCCC
ACAGTTTTAATCCTAGATGAACCTACCGTTGGAATTGATCCATCCTTGAGG
AGAAAAATCTGGCAAGAGCTAATTAATATTAAGGATGAAGGACGTTCTAT
CTTTATTACAACCCACGTTATGGATGAAGCAGAATTAACAAGTAAGGTTGC

ACTACTATTACGTGGAAACATTATTGCCTTTGATACTCCATTACATTTAAA
AAAACAATTTAATGTGAGTACTATTGAGGAAGTTTTCTTAAAAGCTGAAGG
AGAATAA

5 (SEQ ID NO: 220)
MKKVIDLKKLQKAYASETVLNNINLEVFKGEIIGLIGPSGAGKSTLIKTMLGME
KADKGTALVLDTQMPDRNILNQIGYMAQSDALHESLTGLENLLFFGKMKG IQ
KTELKQQITHISKVVDLENQLDKFVSGYSEGMKRRLSLAIALLLGNPTVLILDEP
TVGIDPSLRRKIWQELINIKDEGRSIFITTHVMDEAELTSKVALLLRGNIIAFDTP
10 LHLKKQFNVSTIEEVFLKAEGE*

Sequence description:

15 A] Length: 717 bp - 239 aa (Possible full-length sequence)
B] This gene sequence was not identified using the LEEP system. It was
identified upstream of the ID-130 gene which was identified by LEEP, during
cloning and sequence analysis of the full-length ID-130 gene sequence. ATG
start codon is preceded by a possible
20 Shine-Dalgarno. No obvious potential leader
peptide sequence

ID-188

25 Clone 3-83b (ID-144b)

(SEQ ID NO: 221)
ATGGTACAAATGATACATGATATGATTA AAAACAATTGAGCATT TTTGCTGAG
30 ACACAAGCTGATTTTCCAGTGTATGATATTTTAGGGGAAGTCCATACTTAT
GGACA ACTTAAAGTAGACTCTGACTCTCTAGCTGCTCATATTGATAGCCTA
GGCCTTGTTGAAAAATCACCTGTCTTAGTATTCGGTGGTCAAGAATATGAA
ATGTTGGCGACATTTGTTGCTTTAACAAAGTCAGGGGCATGCTTATATACCG
GTTGACCAACACTCTGCTTTGGATAGAATACAGGCTATTATGACAGTTGCT
35 CAACCAAGCCTTATCATTTC AATTGGTGAATTCCTCTTGAAGTTGATAAT
GTCCCAATCCTAGACGTTTCTCAAGTTTCAGCTATTTTGAAGAAAAGACT
CCTTATGAGGTAACACATTCTGTTAAAGGTGATGATAATTACTATATTATT
TTC ACTTCAGGGACTACTGGTTTACCAAAAGGTGTGCAAATTCACATGAC
AATTTATTGAGCTTACAAATTGGATGATTTCTGATGATGAGFTTTCAGTTC
40 CTGAAAGACCGCAAATGTTGGCTCAACCC

(SEQ ID NO: 222)
MVQMIHDMIKTIEHFAETQADFPVYDILGEVHTY GQLKVDSDSLAAHIDSLGL
VEKSPVLVFGGQEYEMLATFVALTKSGHAYIPVDQHSALDRIQAIMTVAQPSL

IISIGEFPLEVDNVPILDVSQVSAIFEEKTPYEVTHSVKGDDNYIIFTSGTTGLP
KGVQISHDNLLSFTNWMISDDEFSVPERPQMLAQP

5 Sequence description:

- A] Length: 592 bp - 197 aa (partial sequence)
B] This gene sequence was not identified using the LEEP system. It was
10 identified downstream of the ID-144 gene which was identified by LEEP,
during cloning and sequence analysis of the full-length ID-144 gene sequence.
Putative ATG start codon is preceded by a
typical Shine-Dalgarno sequence. No obvious
leader peptide sequence
15 This orf is not in frame with nuc

ID-189

20 Clone 3-86b (ID-145b)

(SEQ ID NO: 223)
ATGGAAAATCATCGTTATGAAGATGAAGGTAAATTCCAGCGTAAGATGAC
CAGTCGTCATCTCTTTATGTTATCGCTAGGTGGTGTATCGGGACTGGGCTT
25 TTCTTGAGTTCAGGTTATACCATTGCACAGGCTGGTCCGCTTGGAGCTGTG
CTGTCTTATTTGATTGGTGCCGTTGTGGTTTATTTGGTCATGCTATCACTTG
GGGAATTGGCGGTTGCCATGCCGGTGACGGGGTCATTCCACACTTATGCCA
CTAAGTTTATCAGTCCTGGAACAGGTTTTACTGTTGCTTGGCTATATTGGAT
TTGTTGGACGGTCGCCTTGGGGACTGAATTTTATAGGTGCTGCCATGCTGAT
30 GCAGCGCTGGTTCCCAAATGTGCCGGCTTGGGCATTTGCTTCCTTTTTTGCC
CTTGTGATTTTTGGTTTAAATGCTCTTAGCGTACGCTTTTTTGCAGAAGCAG
AGTCTTTCTTCTCAAGTATTAAGGTTATTGCTATCATTATCTTTATTATCTTG
GGCTTAGGTGCTATGTTTGGTCTAGTTTTCCTTTGAAGGTCAGCACAAAGGCT
ATTCTCTTCACTCATCTGACTGCCAATGGTGCCTTTCCAAATGGTATCGTTG
35 CAGTTGTCTCAGTCATGTTGGCTGTAACTATGCCTTCTCTGGTACTGAGTT
AATTGGTATTGCGGCTGGTGAAACGGATAATCCCAAAGAAGCTGTACCAA
GGGCTATTAAAACGACAATCGGTCGCTTGGTTGTTTTCTTTGTACTGACAA
TTGTTGTCCTAGCTTCGCTATTGCCAATGAAAGAGGCAGGCGTATCCACAG
CACCATTTCGTTGATGTCCTTGACAAGATGGGAATCCCTTTACGGGCGGATA
40 TCATGAACTTCGTTATCTTGACAGCCATCCTGTCTGCTGGTAACTCAGGTCT
CTACGCATCAAGCCGTATGCTCTGGTCCCTTGCCAATGAAGGTATGTTGTC
AAAATCTGTTGTGAAAATCAATAAACACGGTGTCCCAATGCGTGCTCTTCT
CTTGTCAATGGCAGGAGCAGTGCTGTCGCTCTTTTCAAGTATTTACGCTGC

AGACACAGTTTATCTAGCCTTGGTTTCAATCGCGGGCTTTGCTGTTGTTGTC
GTATGGCTAGCCATTCCAGTCGCACAAATCAATTTCCGCAAGGAATTC

(SEQ ID NO: 224)

5 MENHRYEDEGKFQRKMTSRHLFMLS LGGVIGTGLFLSSGYTIAQAGPLGAVL
SYLIGAVVVYLVMLSLGELAVAMPVTGSFHTYATKFISPGTGFTVAWLYWIC
WTVALGTEFLGAAMLQMQRWFPNVPAAWAFASFFALVIFGLNALSVRFFAEAES
FFSSIKVIAIIIFILGLGAMFGLVSFEGQHKAILFTHLTANGAFPNGIVAVVSVM
10 LAVNYAFSGTELIGIAAGETDNPKEAVPRAIKTTIGRLVVFFVLTIVVLASLLPM
KEAGVSTAPFVDVFDKMGIPFTADIMNFVILTAILSAGNSGLYASSRMLWSLA
NEGMLSKSVVKINKHGVPMRALLSMAGAVLSLFSSIYAADTVYLALVSIAGF
AVVVVWLAIPVAQINFRKEF

15 Sequence description:

A] Length: 1126 bp - 393 aa (partial gene
sequence)
20 B] This gene sequence was not identified using the LEEP system. It was
identified downstream of the ID-145 gene which was identified by LEEP,
during cloning and sequence analysis of the full-length ID-145 gene sequence.
Putative ATG start codon is preceded by a
typical Shine-Dalgarno sequence. Possesses a
25 possible leader peptide sequence.

ID-190

30 Clone 3-94b

(SEQ ID NO: 225)

TCAGAAAATGCAGAGGCAGCAACGGTTGCCACAACTTGGTTACCAAAGG
AGCTAATGTCATTATCGGACCAGCAACATCGGGTGCAGCTGCATCTTCAAC
35 TCCAAAAGTAAATGCAGCAGCAGTTCCAATGATTGCACCTGCTGCGACAC
AAGACAATTTAGTCTATGGTTCTGATGGAAAAACCTTAAATCAGTATTTCT
TCCGAGCTACTTTTGTCGATAATTATCAAGGAAAGCTATTGTCTCAGTATG
CTACAGACAACCTTAAAGCTAAAAAAGTTGTTCTATTTTATGATAATTCAT
CAGATTACTCAAAGGGGGTAGCAAAATCATTTAAGGAAAGTTATAGTGGA
40 AAAATTGTTGATAGTATGACATTCTCCGCTGGTGATACTGATTTCCAAGCG
TCATTGACTAAGTTGAAAGGGGAAAGAATATGATGCTATTGTGATGCCAGG
TTACTATACCGAGACAGGATTAATAGTTAAGCAAGCGCGTGATTTAGGTAT
CTCTAAACCGGTTCTTGGGCCTGATGGTTTTGATAGTCCGAAATTTGTGCA

ATCGGCAACACCTGTAGGAGCTTCAAACGTTTATTATTTGACAGGTTTCAC
TACACAAGGATCAACCAAAGCTAAAGCT

(SEQ ID NO: 226)

5 SENAEAAATVATNLVTKGANVIIGPATSGAAASSTPKVNAAAVPMIAPAATQD
NLVYGSDGKTLNQYFFRATFVDNYQGKLLSQYATDNLKAKKVVLFDNSSD
YSKGVAKSFKESYSGKIVDSMTFSAGDTDFQASLTKLKGKEYDAIVMPGYT
ETGLIVKQARDLGISKPVLGPDGFDSPKFVQSATPVGASNVYYLTGFTTQGST
KAKA

10

Sequence description

15

A] Length: 637 bp - 231 aa (partial sequence)

B] This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-149 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-149 gene sequence. N- and C-termini have yet to be determined

20

ID-191

Clone 2-c94b (ID-153b)

25

(SEQ ID NO: 227)

TTGGGACTTAAAGACCATGCTTTAGTCTATCCATTTTCATTATCTGGGGGGG
CAAAAGCAACGTGTCGCACTAGCTCGTGCGATGATGATTGATCCACAGATT
ATTGGTTATGATGAGCCAACTAGCGCTCTTGATCCAGAGTTGCGTCAAGAA
30 GTAGAAAAACTAATTTTACAAAATAGAGAAACAGGTATGACACAAATTGT
AGTAACACATGATCTTCAATTTGCTGAAAGTATATCTGATACGATTCTCAA
AATTAATCCTAAGTAG

30

(SEQ ID NO: 228)

35

MGLKDHALVYPFSLSGGQKQRVALARAMMIDPQIIGYDEPTSAIDPELRQEV
EKLILQNRETGMTQIVVTHDLQFAESISDTILKINPK*

Sequence description

40

A] Length: 270 bp - 90 aa (partial sequence)

B] This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-153 gene which was identified by LEEP, during cloning and sequence analysis of the ID-153 gene sequence.
N-terminus has yet to be determined

5

ID-192

Clone 2-c1b (ID-155b)

10

(SEQ ID NO: 229)

ATGACTAATATCTCAGATGTTCCAAAAGCTATTAGAACACAGGCACAGTAT
GTTCTCTTGGGAATGAGAGTTATGGATCAGTCGGTATTACCGAAAACATAT
AATTCAAAAGAACCTTATTTGAAACCAGATATGATTTATATTCATGATAGA
15 AGACAAGAGACAATGCTTAAAATCACTCAAGAAATAGAAATGGAGCATTG
A

(SEQ ID NO: 230)

20

MTNISDVPKAIRTQAQYVLLGMRVMDQSVLPKTYNSKEPYLKPDMIYIHDRR
QETMLKITQEIEMEH*

Sequence description

25

A] Length: 204 bp - 68 aa (partial sequence)

B] This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-155 gene which was identified by LEEP, during cloning and sequence analysis of the ID-155 gene sequence.

30

ATG start codon is preceded by a potential typical Shine-Dalgarno sequence.
Has a
typical leader peptide. N-terminus has yet to be
verified

35

ID-193

Clone 2-54altb (ID-172b)

40

(SEQ ID NO: 231)

AAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCTTGGGGAATATAAATT
TGGATTTTCATGACGATGTAAAGCCAATTTATTCTACGGGAAAAGGTCTAAA
TGAGGCTGTTATTCGTGAGTTATCTGCAGCTAAGGGTGAACCTGAGTGGAT
GTTGGACTTTCGTCTAAAATCCTTGGAACGTTTAATAAAAATGCCGATGCA

5 GACCTGGGGAGCAGATTTATCAGATATTGATTTTGATGATATTATTTATTA
 TCAAAAAGCATCTGATAAACCTGCGCGTGATTGGGATGATGTTCCAGAAA
 AAATCAAAGAACTTTTGAAAGAATTGGGATTCCAGAAGCTGAAAGAGCC
 TATCTTGCAGGAGCATCAGCACAAATATGAATCAGAAGTAGTTTATCACAAT
 10 ATGAAAGAAGAATATGATAAGCTGGGTATTGTTTTTACGGATACTGACTCT
 GCACTTAAAGAGTACCCAGAGCTATTCAAAAAATATTTTGCTAAACTTGTC
 CCTCCAACAGATAATAAATTAGCTGCTCTGAACTCTGCTGTATGGTCAGGT
 GGAACATTTATTTATGTTTCCTAAAGGTGTTAAGGTGGATATTCCACTTCAA
 ACTTACTTCCGTATTAATAATGAAAATACTGGACAATTTGAACGTA CTCTC
 ATTATTGTTGATGAGGGAGCAAGTGTTCACTATGTTGAAGGTTGTACCGCC
 CCAACTTATTCTTCAAATAGTTTACATGCAGCTATAGTTGAAATTTTTGCAC
 TTGATGGAGCTTATATGCGCTATACGACTATTCAAAATTGGTCCGATAATG
 TCTATAATTTAGTGACAAAACGTGCTACCGCTAAAAAAGATGCAACAGTT
 15 GAGTGGATAGATGGAAATCTAGGAGCTAAAACAACAATGAAATACCCATC
 GGTTTACCTTGATGGTGAAGGAGCACGTGGCACGATGTTGTCTATTGCTTT
 TGCAAACAAAGGACAACACCAAGATACGGGTGCAAAGATGATTCATAATG
 CCCCCATACTAGTTCATCCATTGTCTCTAAATCAATTGCTAAGGGTGGGG
 GAAAAGTTGATTATCGAGGTCAAGTGACATTTAATAAAGATTCCAAAAAA
 TCAGTGTCACATATAGAATGTGACACCATATTGATGGATGATATTTCAAAA
 20 TCAGATACCATAACGTTTAAATGAGATTCATAATTCACAGGTTGCTTTAGAG
 CATGAAGCAAAGGTGTCTAAGATTTCTGAAGAGCAACTGTACTACTTGATG
 AGTCGAGGTTTATCTGAAGCTGAAGCAACAGAAATGATTGTTATGGGGTTT
 GTTGAGCCCTTTACGAAAGAATTACCAATGGAATATGCGGTAGAGTTAAA
 TCGTTTAATTCCTATGAAATGGAAGGTTCAAGTTGGTTAA

25

(SEQ ID NO: 232)

MHACRSTLEDLGEYKFGFHDDVKPIYSTGKGLNEAVIRELSAAKGEPEWMLD
 FRLKSLETFNKMPMQTWGADLSIDFDDIIYYQKASDKPARDWDDVPEKIKE
 TFERIGIPEAERAYLAGASAQYESEVVYHNMKEEYDKLGIVFTDSDSALKEYP
 30 ELFKKYFAKLVPPTDNKLAALNSAVWSGGTFIYVPKGVKVDIPLQTYFRINNE
 NTGQFERTLIIVDEGASVHYVEGCTAPTYSSNSLHAAIVEIFALDGAYMRYTTI
 QNWSDNVYNLVTKRATAKKDATVEWIDGNLGAKTTMKYPSVYLDGEGARG
 TMLSIAFANKGQHQDTGAKMIHNAPHTSSSIVSKSIKGGGKVDYRGQVTFN
 KDSKKS VSHIECDTILMDDISKSDTIPFNEIHNSQVALEHEAKVSKISEEQLYYL
 35 MSRGLSEAEATEMIVMGFVEPFTKELPMEYAVELNRLISYEMEGSVG*

Sequence description:

40

A] Length: 1411 bp - 469 aa (Possible full-length gene)

B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-72 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-72 gene sequence. No obvious Shine Dalgarno sequence upstream of

TTG start codon insufficient sequence data). N terminus needs verification.

5 ID-194

Clone 3-1b (ID-81b)

(SEQ ID NO: 233)

10 ATGATAGAATTCTTTTCTAATATCAGAACAGAGATTCCGCAGATGCCTTTA
CTTATCCATAGTTTGGATTTTATCTGTCTTACCTTTTCTGATGTGGCTGACTTT
GGTTAATAGAGATAAGCCTTTGTATAAAACTATTTGGAGTATCCTTTTAGG
ACTTCAGTTAATTACGATTTATACTTGGTTTTTCTGGGCAAATGTCCTTTA
15 TCTGAAAGTCTTCCCCTTTACCATTGTCTGAATAGGCATGTTTGTCGGTCTCT
TA

(SEQ ID NO: 234)

MIEFFSNIRTEIPQMPLLIHSLILSVLPFLMWLTLVNRDKPLYKTIWSILLGLQLI
TIYTWFFWAKLPLSESLPLYHCRIGMFVGLL
20

Sequence description

A) Length: 261 bp - 87 aa (partial gene sequence)

25 B) This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-81 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-81 gene sequence. Sequence Characteristics: Possesses a potential leader peptide sequence Orf is preceded by a potential Shine-Dalgarno sequence.

30

ID-195

Clone RS-55b

35 (SEQ ID NO: 235)

AAGCTTGTGCAAAGTATTAAAGAGATAGGATTAGCTAATGCGCATTTATTA
GCTGTTGCTCCGACAGGGTCAATCAGTTATCTTTCTTCTTGTACTCCGAGCC
TTCAACCGGTTGTATCACCTGTCTGAAGTACGCAAGGAAGGAGCACTGGGG
AGGGTTTATGTAGCTGCTTATAAGATTGATGCAGATAATTATGTCTAETAC
40 AAAAAAGGAGCTTATGAAGTGGGATCTGAGGCGATTATCAATATTGCAGC
TGCCGCTCAAAAACACATTGATCAAGCTATTTTCGTTAACGCTTTTCATGAC
AGATCAAGCAACTACGCGAGATTTAAATAAAGCCTATATTCAAGCATTTA
AACAAAAATGTGCCTCTATTTATTATGTACGAGTGAGACAGGACATCCTAG

AAGGTAGCGAGAGTTATGATGATATGCTGGATGATTTCACCTTCATCGGACT
TAGAAGACTGTCAATCCTGCATGATTAA

(SEQ ID NO: 236)

5 KLVQSIKEIGLANAHLAVAPTGSISYLSSCTPSLQPVVSPVEVRKEGALGRVY
VAAYKIDADNYVYYKKGAYEVGSEAIINIAAAAQKHIDQAISLTLFMTDQATT
RDLNKAYIQAFKQKCASIYYVRVRQDILEGSESYDDMLDDFTSSDLEDCQSC
MI*

10

Sequence description:

A] Length 486 bp - 162 aa (Partial sequence)

15 B] This gene sequence was not identified using the LEEP system. It was identified
upstream of the ID-87 gene which was identified by LEEP, during cloning and
sequence analysis of the full-length ID-87 gene sequence. N-terminus to be
determined.

20

ID-196

Clone RS-59(ID-90b)

(SEQ ID NO: 237)

25 GTGAGGACATATATTACAAACTTGAATGGACATTCAATCACTAGTACAGC
ACAAATAGCTCAAAACATGGTAACAGATATAGCAGTAAGCTTAGGTTTTC
GTGAGCTGGGAATACATTCTTATCCGATTGATACTGATTCTCCTGAGGAAA
TGAGTAAGCGTTTAGATGGAATCTGTTCCGGACTTAGAAAAAATGATATTG
TCATATTTTCAGACACCTACATGGAACACTACAACCTTTTGATGAAAAATTAT
30 TTCACAAATTAAAAATATTTGGTGTAAAGATTGTTATTTTTTATACATGATGT
TGTACCGCTAATGTTTGTATGGAAATTTTTTATTTGATGGATAGAACTATAGC
TTATTATAATGAAGCAGATGTTTAATAGCCCCTAGTCAAGCAATGGTCGAT
AAGCTT

35

(SEQ ID NO: 238)

MRTYITNLNGHSITSTAQIAQNMVTDIAVSLGFRELGIHSPIDTDSPEEMSKRL
DGICSGLRKNDIVIFQTPTWNTTTFDEKLFHKLKIFGVKIVIFIHDTVPLMFDGN
FYLMDRITIAYYNEADVLIAPSQAMVDKL

40

Sequence description:

A] Length: 414 bp - 138 aa(partial gene)

5 B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-90 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-90 gene sequence. No obvious signal peptide, but a possible Shine Dalgarno sequence is present upstream of ATG start codon. C-terminus has yet to be determined.

10 ID-197

Clone RS-59c (ID-90c)

(SEQ ID NO: 239)

15 CATGGAAATGAAGTTGATGATGTTATTAGAAGGGCATTGGAATATAATCAC
CTTATCTTTGCTTTTGATAATACCTGTCATAACAGAGAGTTAGTATTAGATA
GCAATATCATTTCTCACACAACCTGTGAACAATTGATAAATTTAATGAAAA
ATTTATCAGGCTCCATTATGTATTTGCTAGAGCAACAAAGAGAACAAACA
20 AGTAATGAAACAAAAGAGCGTTATAAAGAAATATTAGGAGGGGTATGGAA
ATGCCTAA

(SEQ ID NO: 240)

25 HGNEVDDVIRRAFEYNHLIFAFDNTCHNRELVLDSNIISHTTCEQLINLMKNLS
GSIMYLLEQQREQTSNETKERYKEILGGYGNA*

Sequence description:

30 A] Length: 261 bp - 87 aa(partial gene sequence)
B] This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-90 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-90 gene sequence. N-terminus has yet to be determined

35

ID-198

Clone RS-70b (ID-93b)

(SEQ ID NO: 241)

40 ACATTTTTATATTATGTATTTGAAGACGTAGCCACCCAGTCAAATATGACT
GGGAAGATTTTATAGTATGTCTAAAGAAGAGTTGTCATATTTACCCGTTATT
AAACTTTTAAAGAATCAAGGTGTATACAACGGCTTGATTGGTCTATTCCTC
CTTTATGGGTTATATATTTTACAGAATCAAGAAATTGTAGCTATTTTTTTAA

TCAATGTGTTGCTAGTTGCTGTTTATGGTGCTTTGACAGTTGATAAAAAAA
TCTTATTAAACAGGGTGGTTTACCTATATTAGCTCTTTTAACATTCTTATT
TTAA

5 (SEQ ID NO: 242)
TFLYYVFEDVATQSNMTGKIFSMSKEELSYLPVIKLFKNQGVYNGLIGLFLLY
GLYISQNQEIVAIFLINVLLVAVYGALTVDKKILLKQGGLPILALLTFLF*

10 Sequence description:

A] Length: 312 bp - 104 aa (partial gene sequence)

15 B] This gene sequence was not identified using the LEEP system. It
was identified upstream of the ID-93 gene which was identified by
LEEP, during cloning and sequence analysis of the full-length ID-93
gene sequence.

N-terminus has yet to be determined

20 ID-199

Clone RS-70c (ID-93c)

(SEQ ID NO: 243)
25 ATGAAATTAAGTGTCCTTGATTATGGGCTTATTGATTATGGAAAACTGCA
AGTGATGCAATACAAGAAACGATTCTTTTATCACAAGAGGCGGAGCAACT
AGGCTATCATCAATTTTGGGTGGCTGAACATCACGGTGTTAAGGCATTTCAG
TATTAGCAATCCAGAATTAATGATAATGCATTTGGCTAACCAGACTAAATC
TATCAAAATTGGCTCTGGAGGTATAATGCCTCTGCACTATAGTAGTTTTAA
30 ACTCGCGGAGACTCTCAAGACATTAGAGACATGTCATCCGAATCGAGTAA
GTATTGGTTTAGGAAATTCAGTAGGGACAGTTAAAGTTTCAAATGCACTTC
GTAGCTTACATAAAGCACATGATTACGAAGAGGTACTGGAGGAATTGAAG
TCATGGCTTATTGATGAATCATCCAGTAAGGAACCATTAGTTCAACCGACT
CTTTCTAGCTTCCCAGACTTATATGTGTTGGGGAGTGGTCAAAAATCAGCT
35 TATTTAGCGGCTAAACTTGGCTTAGGCTTTACCTTCGGTGTTTTTCCTTTTA
TGGACAAAGACCCATTGACAGAAGCTAAA

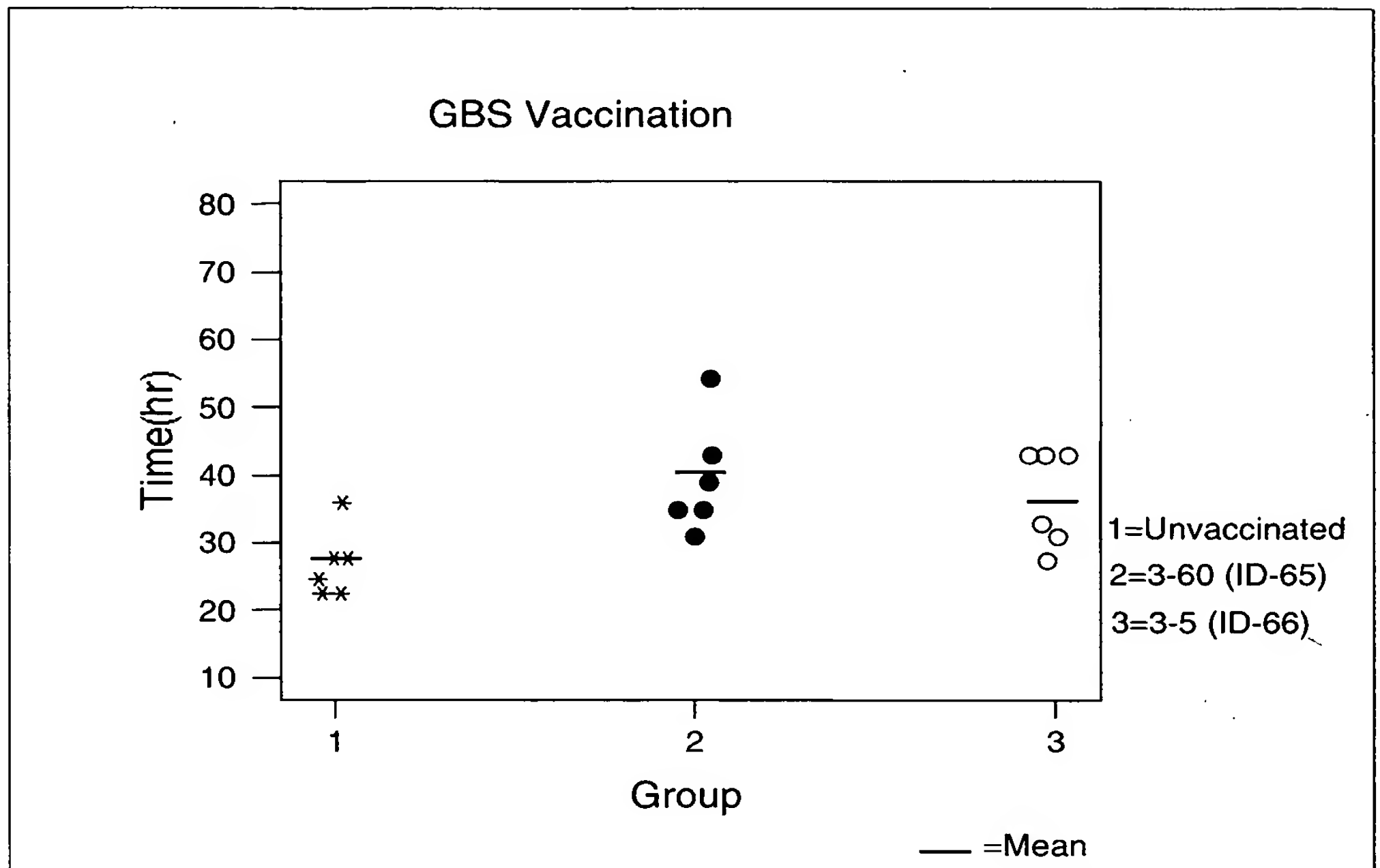
(SEQ ID NO: 244)
40 MKLSVLDYGLIDYGKTASDAIQETILLSQEAEQLGYHQFWVAEHHGVKAFFSIS
NPELMIMHLANQTKSIKIGSGGIMPLHYSSFKLAETLKTLETCHPNRVSIGLGN
SLGTVKVSNALRSLHKAHDYEEVLEELKSWLIDESSSKEPLVQPTLSSFPDLYV
LGSGQKSAYLAAKLGLGFTFGVFPFMDKDPLTEAK

Sequence description:

- 5 A] Length: 588 bp - 196 aa (partial)
B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-93 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-93 gene sequence. No obvious signal peptide, but Shine Dalgarno sequence upstream of the ATG start codon.



Figure 2



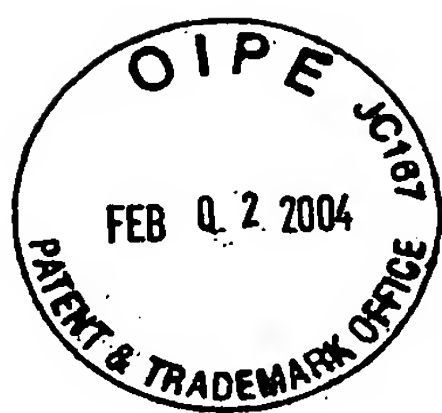


Figure 3

nucS1

Bgl II Eco RV

5'-cgagatctgatatctcacaaacagataacggcgtaaataag -3' (SEQ ID NO.: 245)

nucS2

Bgl II Sma I

5'-gaagatcttccccgggatcacaaacagataacggcgtaaataag -3' (SEQ ID NO.: 246)

nucS3

Bgl II Eco RV

5'-cgagatctgatatccatcacaaacagataacggcgtaaataag -3' (SEQ ID NO.: 247)

nucR

Bam HI

5'-cgggatcccttatggacctgaatcagcgttgto -3' (SEQ ID NO.: 248)

NucSeq

5'-ggatgctttgtttcaggtgtatc -3' (SEQ ID NO.: 249)

PTREP_F

5'-catgatatcggtacctcaagctcatatcattgtccggcaatggtgtgggctttttttgttttagcggataa
caatttcacac -3' (SEQ ID NO.: 250)

PTREP_R

5'-gcggatcccccggggttaattaatgtttaaacactagtcgaagatctcgggaattctctgtgtgaaatt
gttatccgcta -3' (SEQ ID NO.: 251)

pUC_F

5'-cgccagggttttcccagtcacgac -3' (SEQ ID NO.: 252)

V_R

5'-tcagggggggcggagcctatg -3' (SEQ ID NO.: 253)

V₁

5'-tcgtatgttggtggaattgtg -3' (SEQ ID NO.: 254)

V₂

5'-tccggctcgtatgttggtggaattg -3' (SEQ ID NO.: 255)



pTREP-Nuc vectors allow cloning of genomic DNA into each frame with respect to the nuclease gene

5

(i)

```

pTREp1-nuc1 (EcoRV) AAGTATCAGATCT--GATATC--TCACAAACAGATAACGGCGTAAAT Frame=+1
      :               ▲               :
      (SEQ ID NO.:256) :               :
pTREp1-nuc2 (Bsa 1) AAGTATCAGATCTTCCCCCGCA-TCACAAACAGATAACGGCGTAAAT Frame=+2
      :               ▲               :
      (SEQ ID NO.: 257) :               :
pTREp1-nuc3 (EcoRV) AAGTATCAGATCT--GATATCCAACAGATAACGGCGTAAAT Frame=+3
      :               ▲               :
      (SEQ ID NO. 258) :               :
Nuclease Gene          TCACAAACAGATAACGGCGTAAAT
                        (SEQ ID NO.: 259)
Cloning site is indicated by an arrow

```

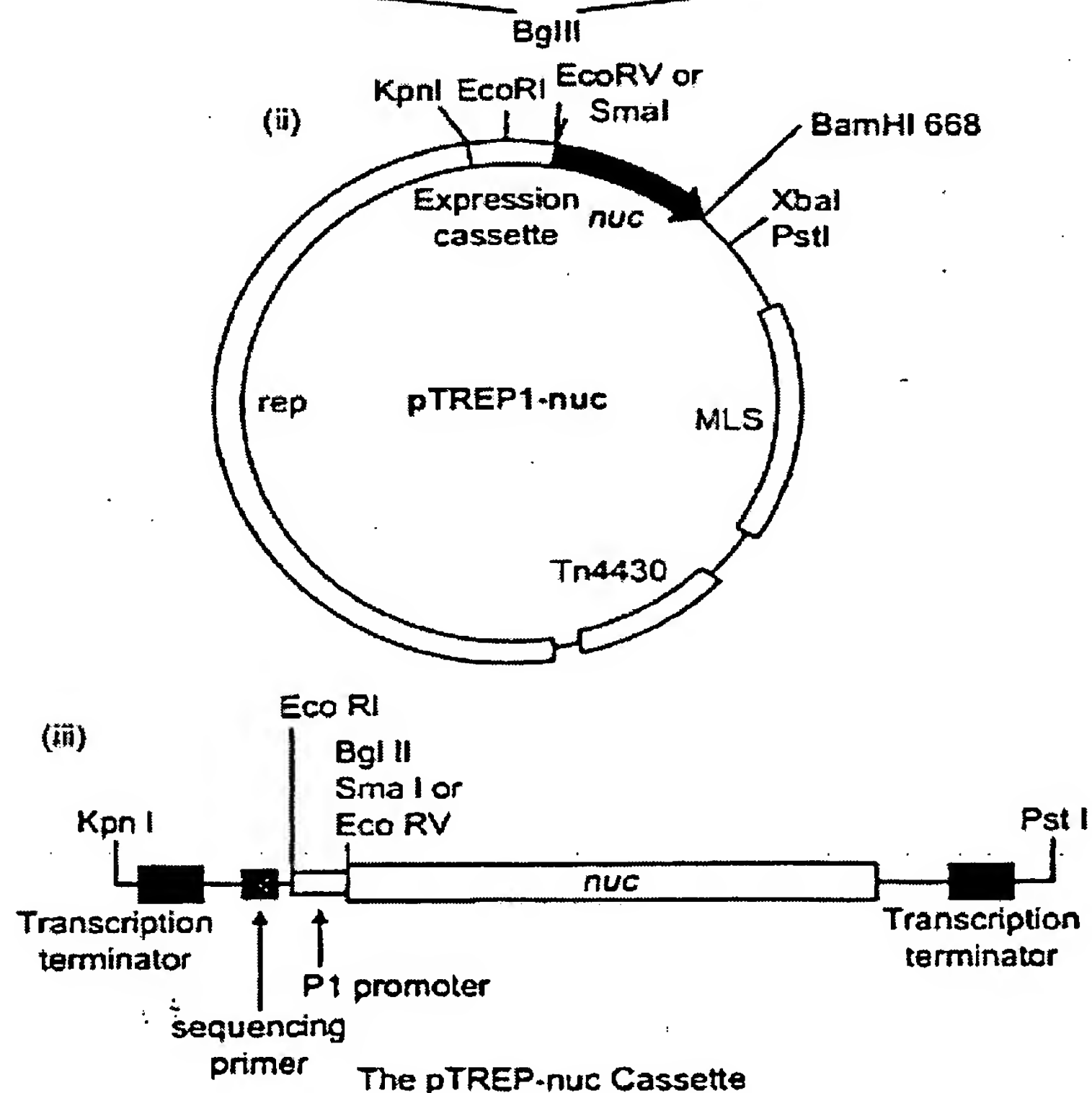


Figure 5

SDS-PAGE analysis of the purified ID-65 and ID-83 protein antigens

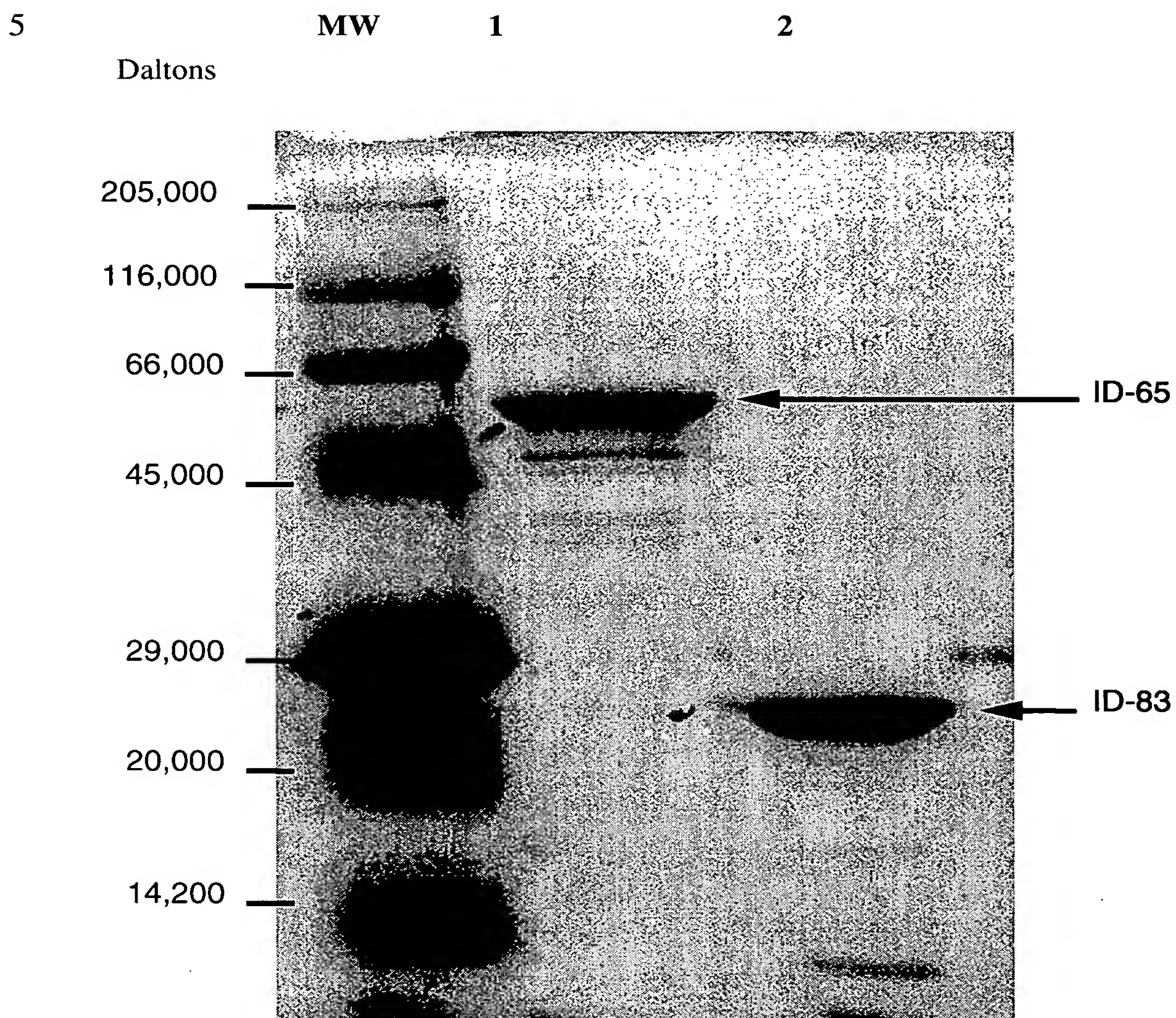




Figure 6

SDS-PAGE analysis of the purified ID-93 antigen

5

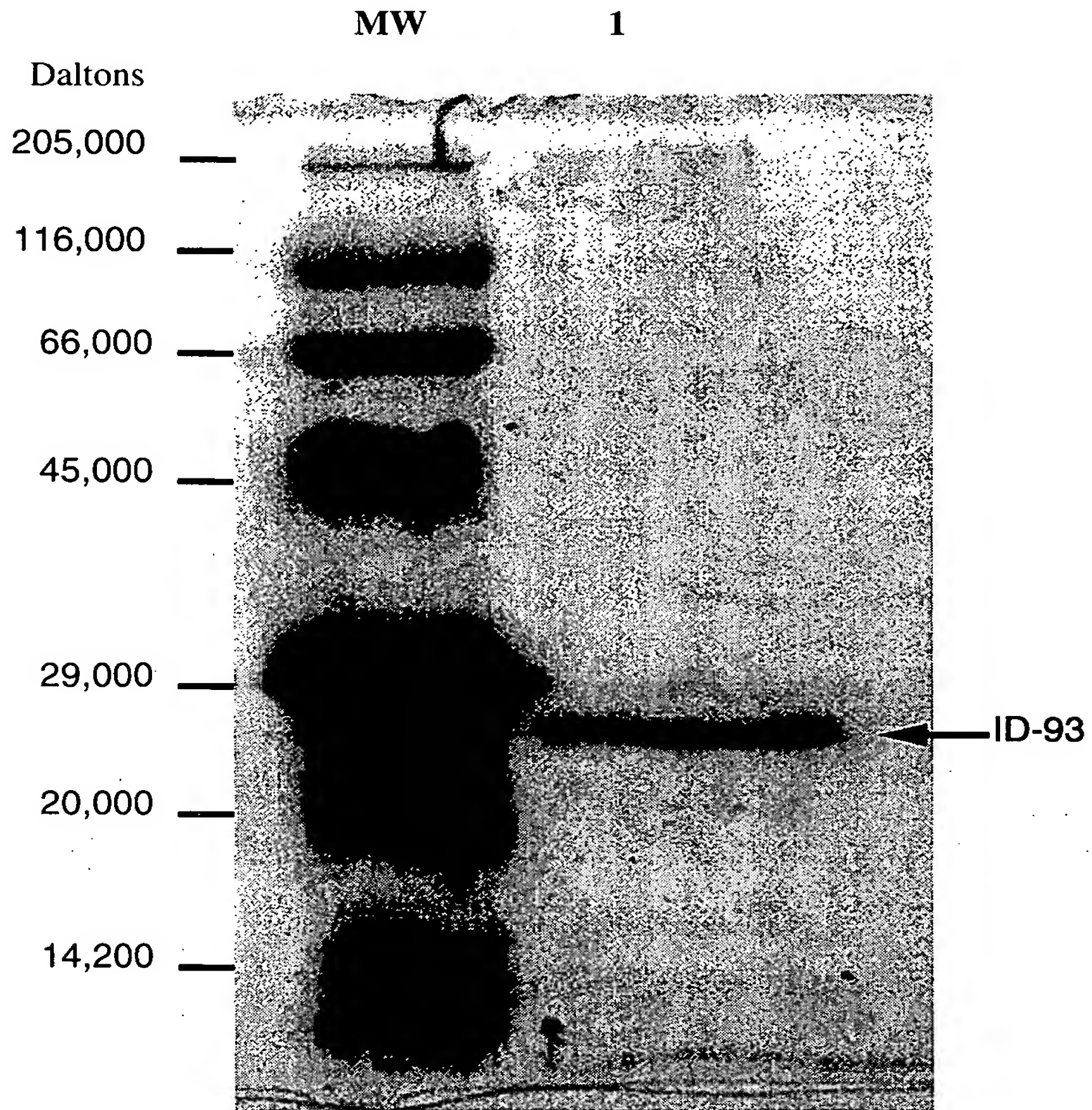
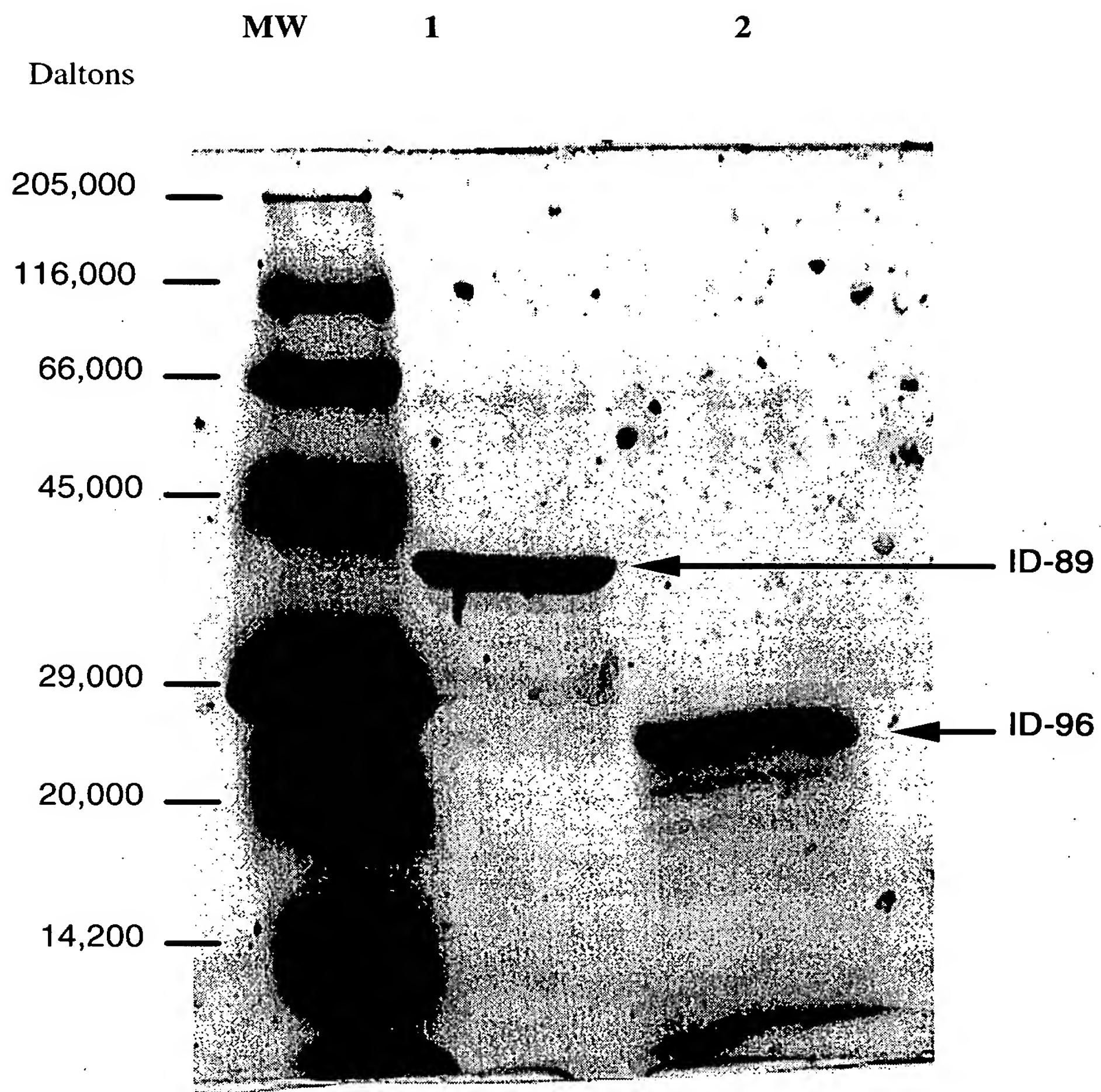




Figure 7

SDS-PAGE analysis of the purified ID-89 and ID-96 protein antigens

5



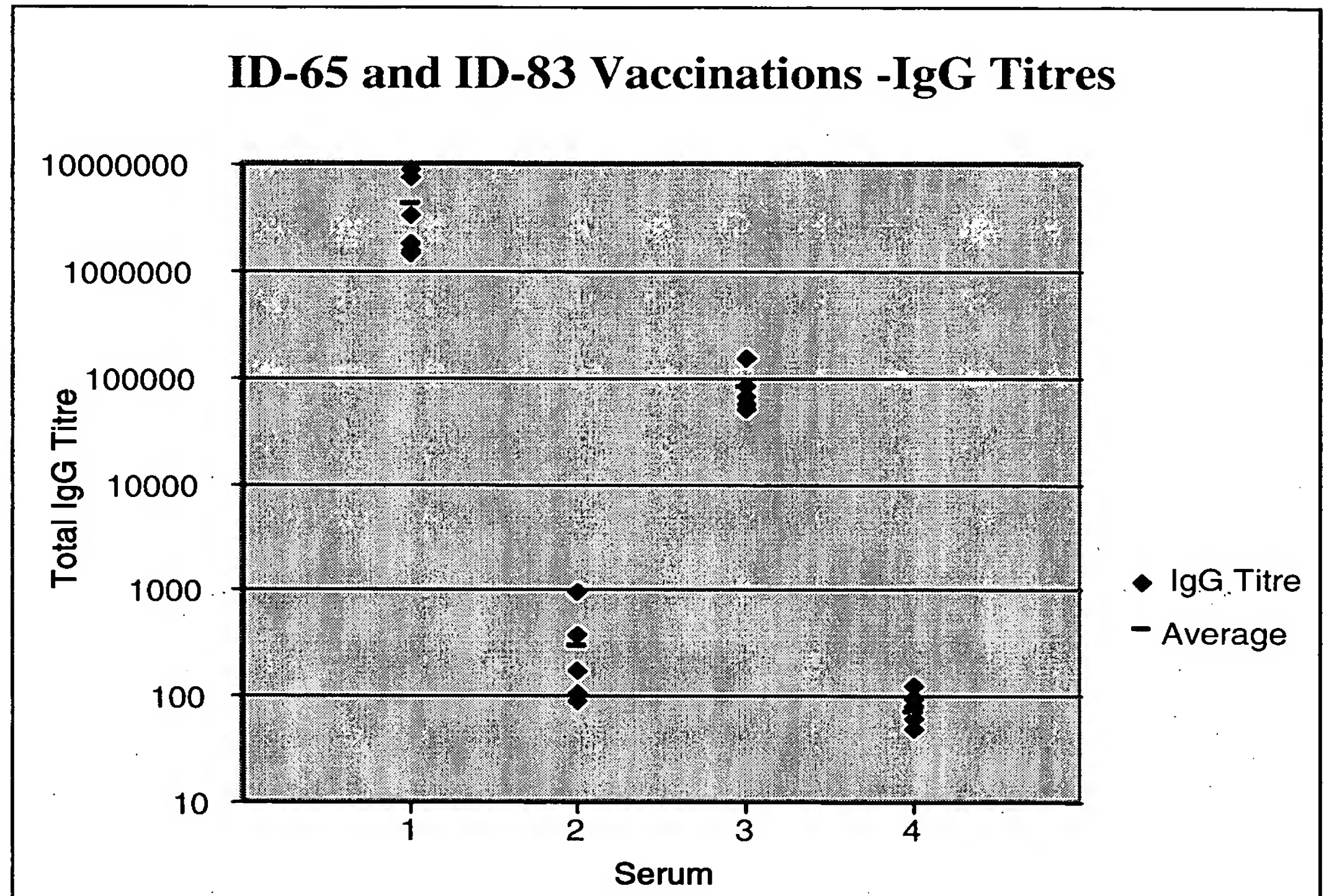
10



Figure 8

IgG Titres against the ID-65 and ID-83 Proteins

5



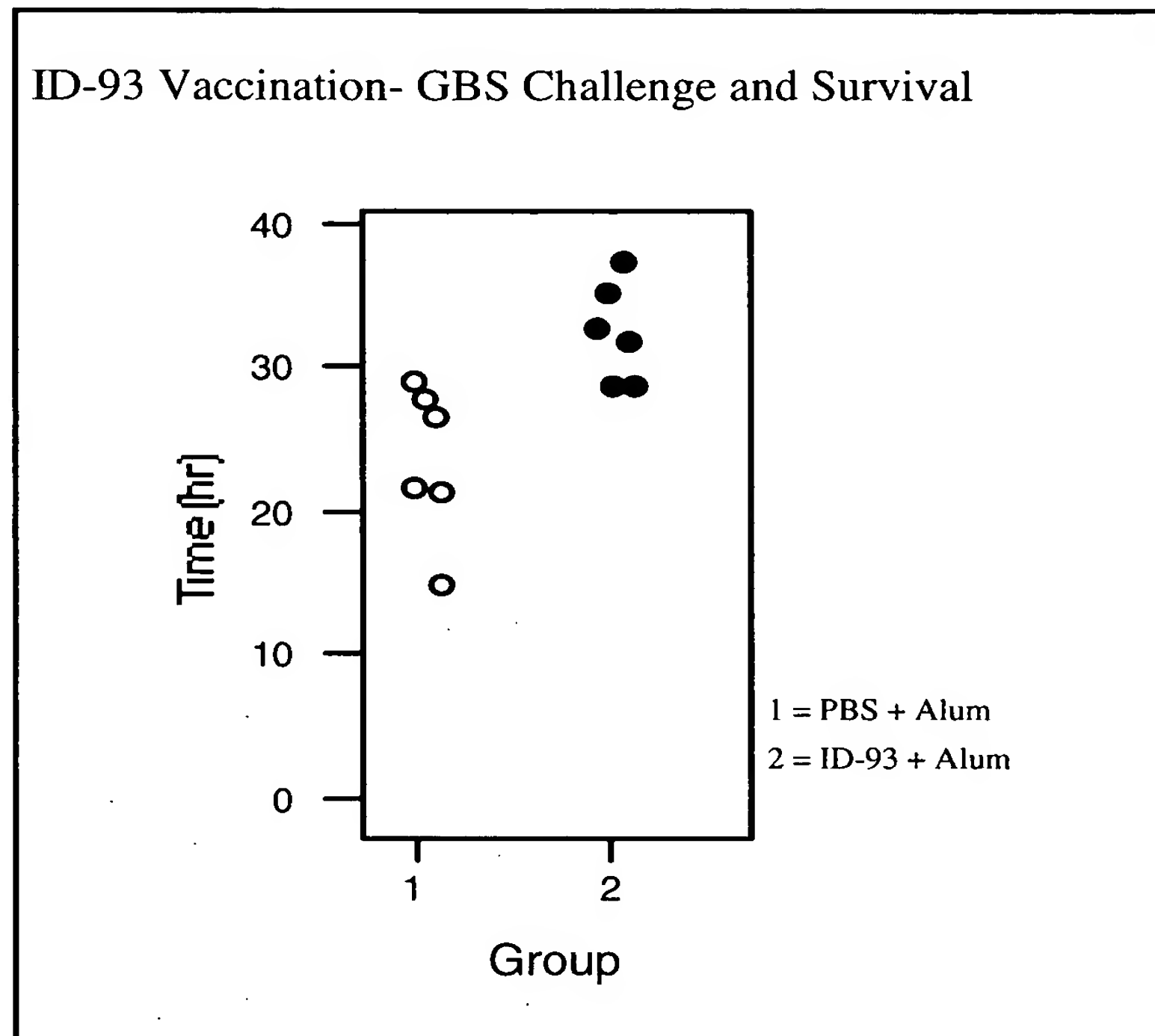
10

15



Figure 9

Survival Data



5

10

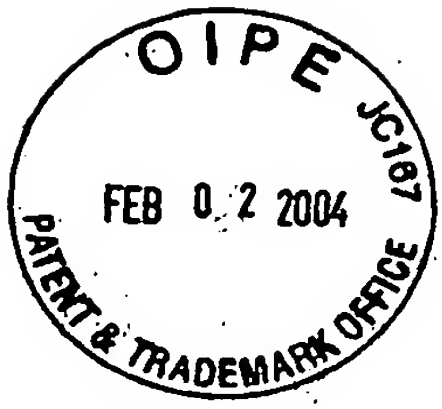


Figure 10

IgG Titres against the ID-93 Protein

5

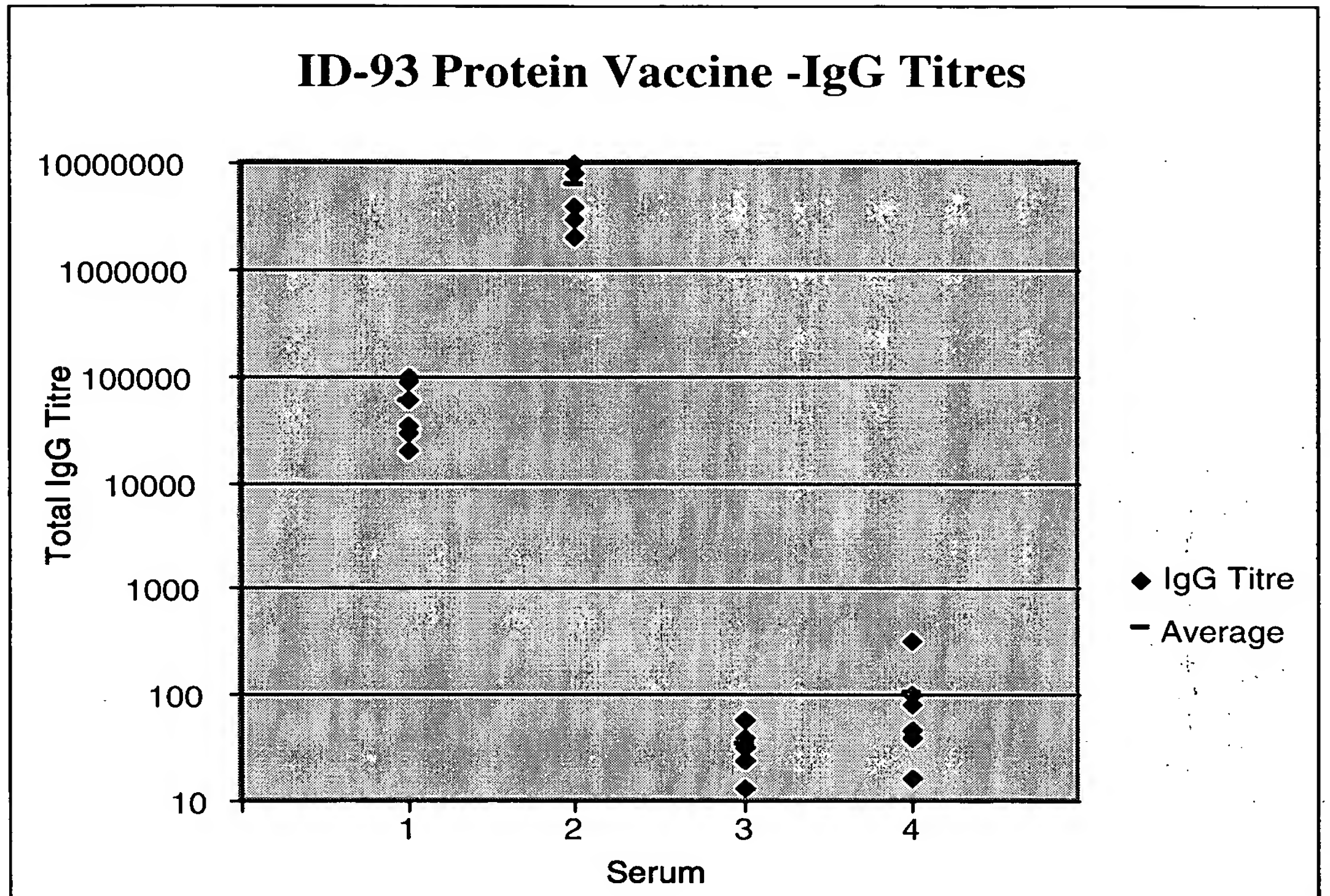




Figure 13 Southern blot analysis – ID-65

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

5

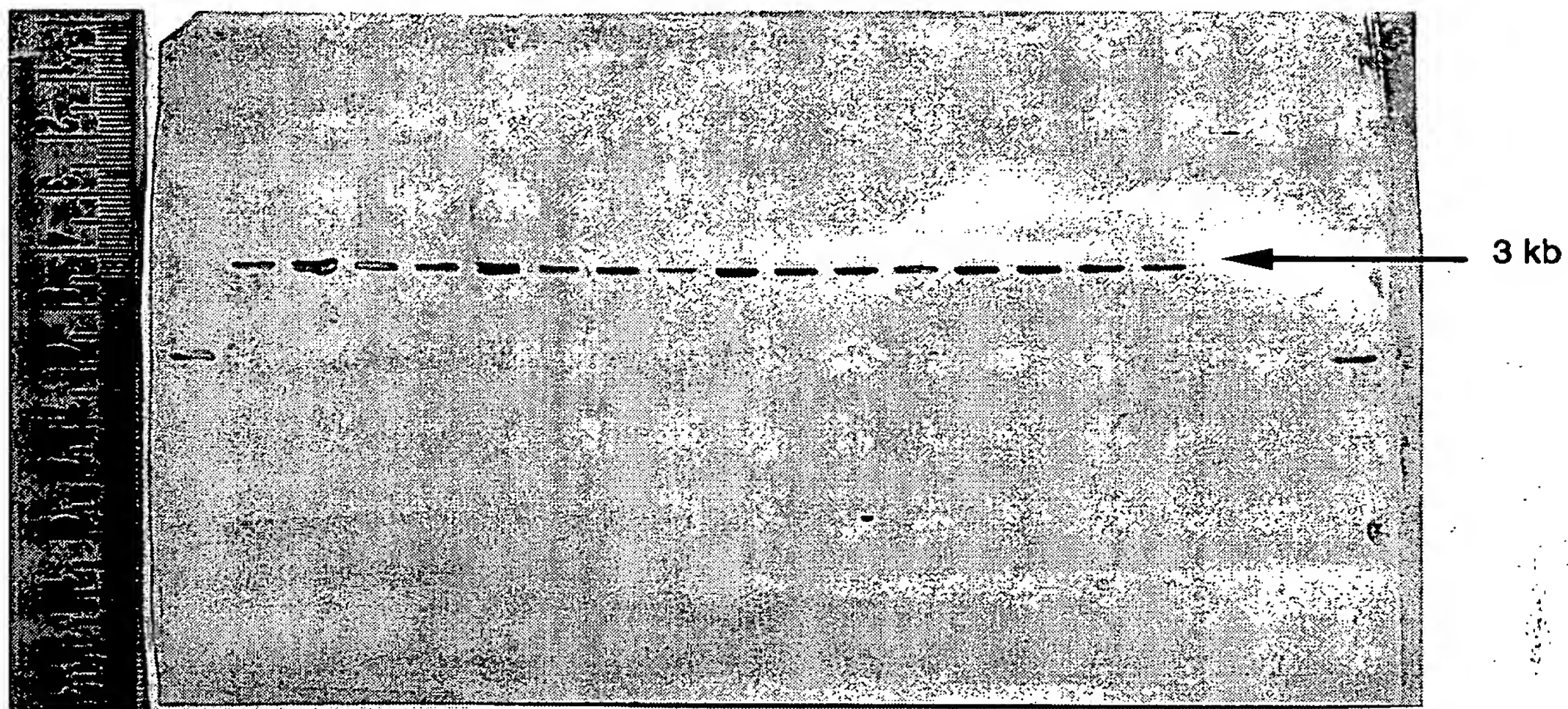
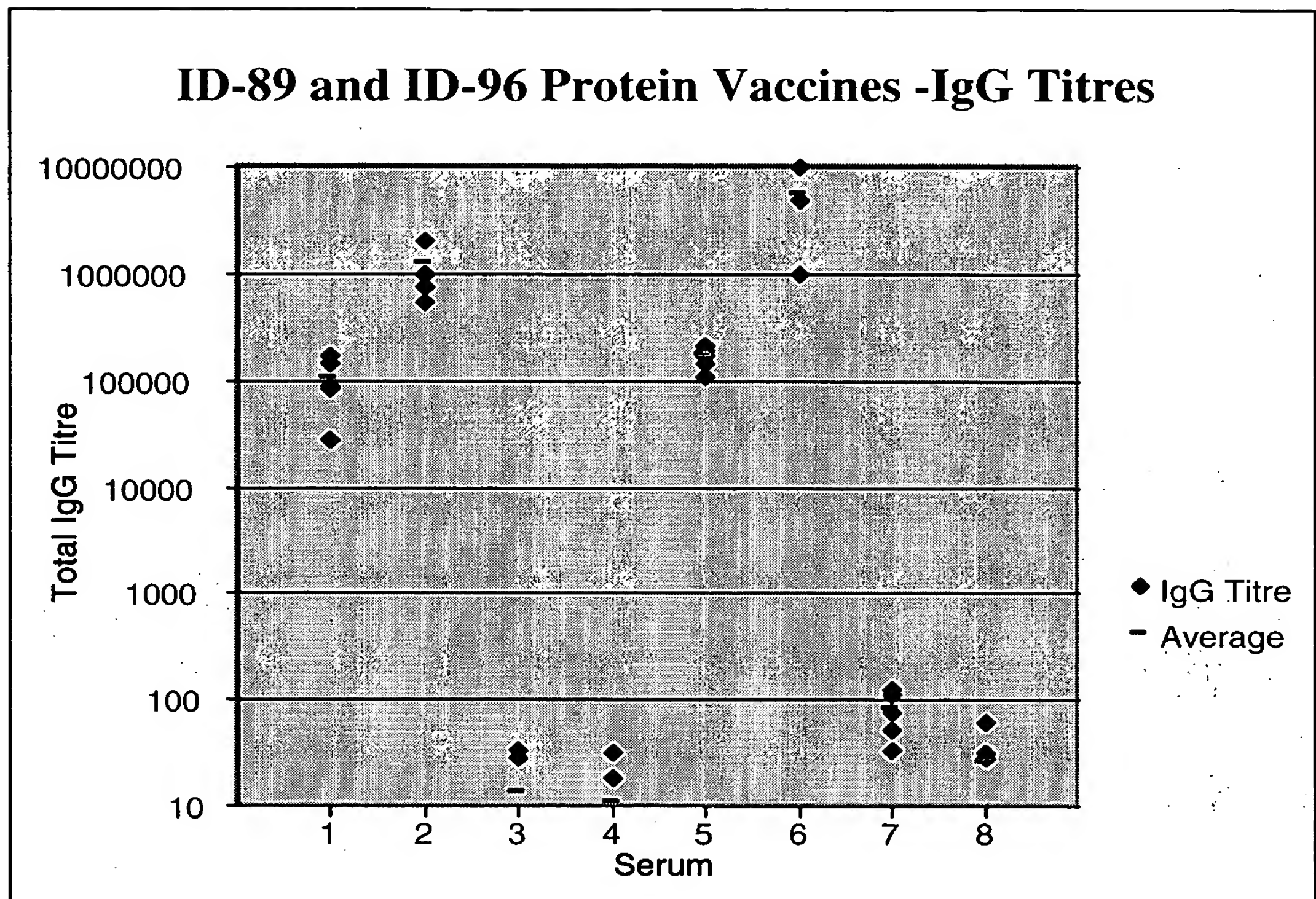




Figure 11

IgG Titres against the ID-89 and ID-96 Proteins



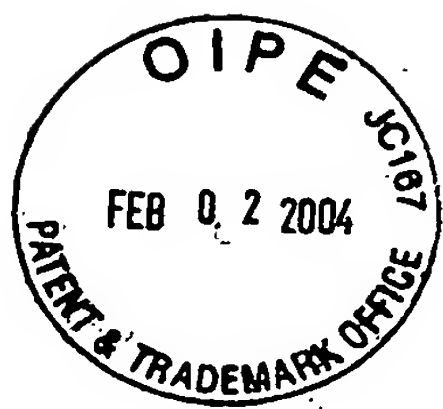
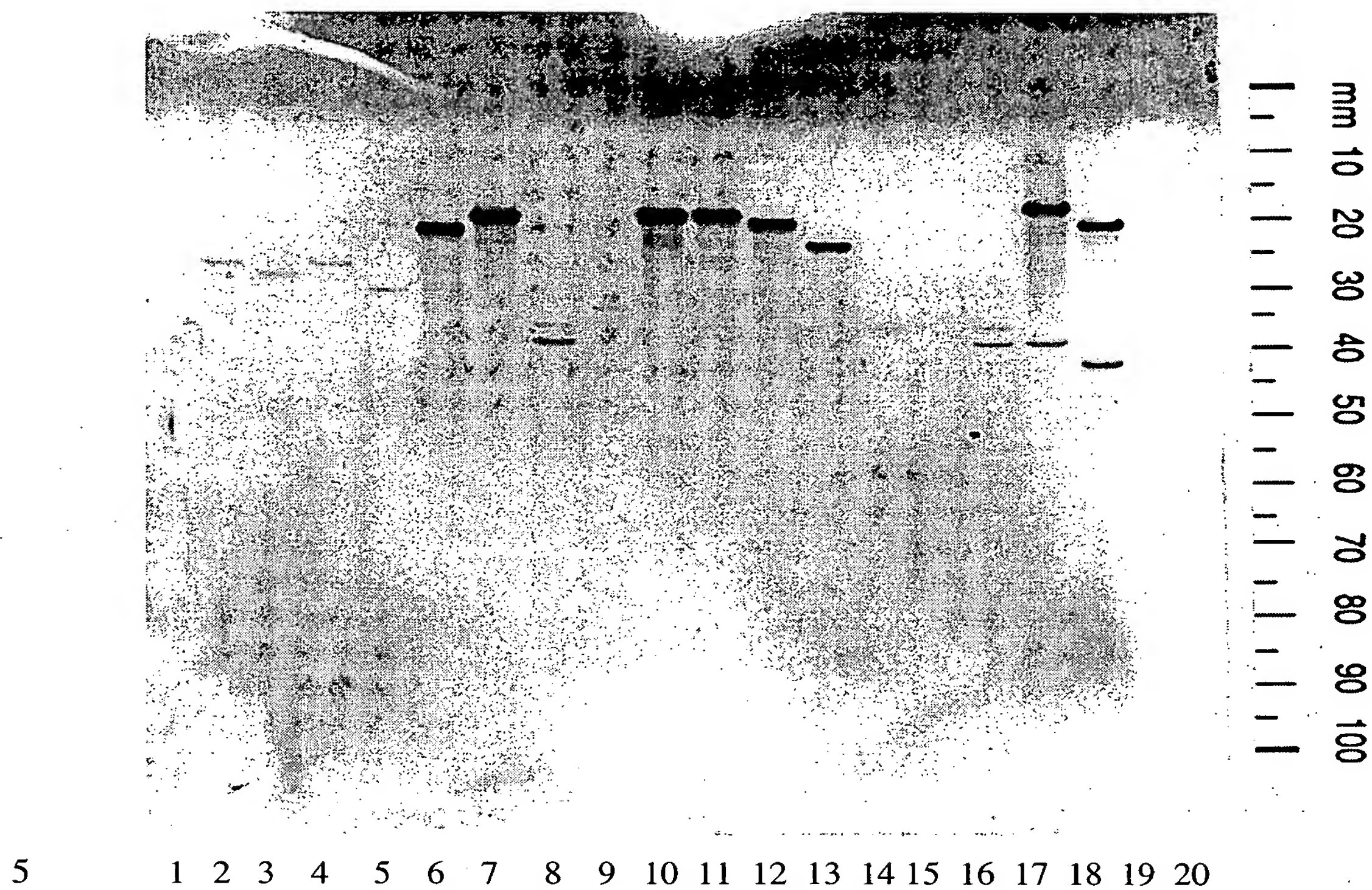


Figure 12

Southern blot analysis - *rib*



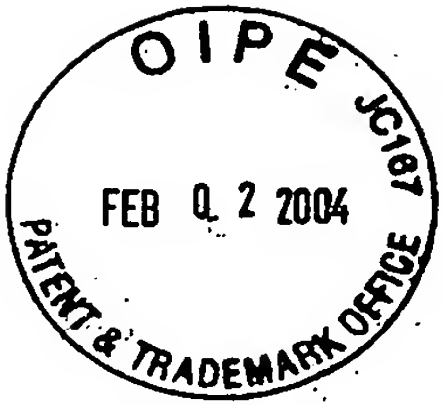


Figure 14

Southern blot analysis – ID-89

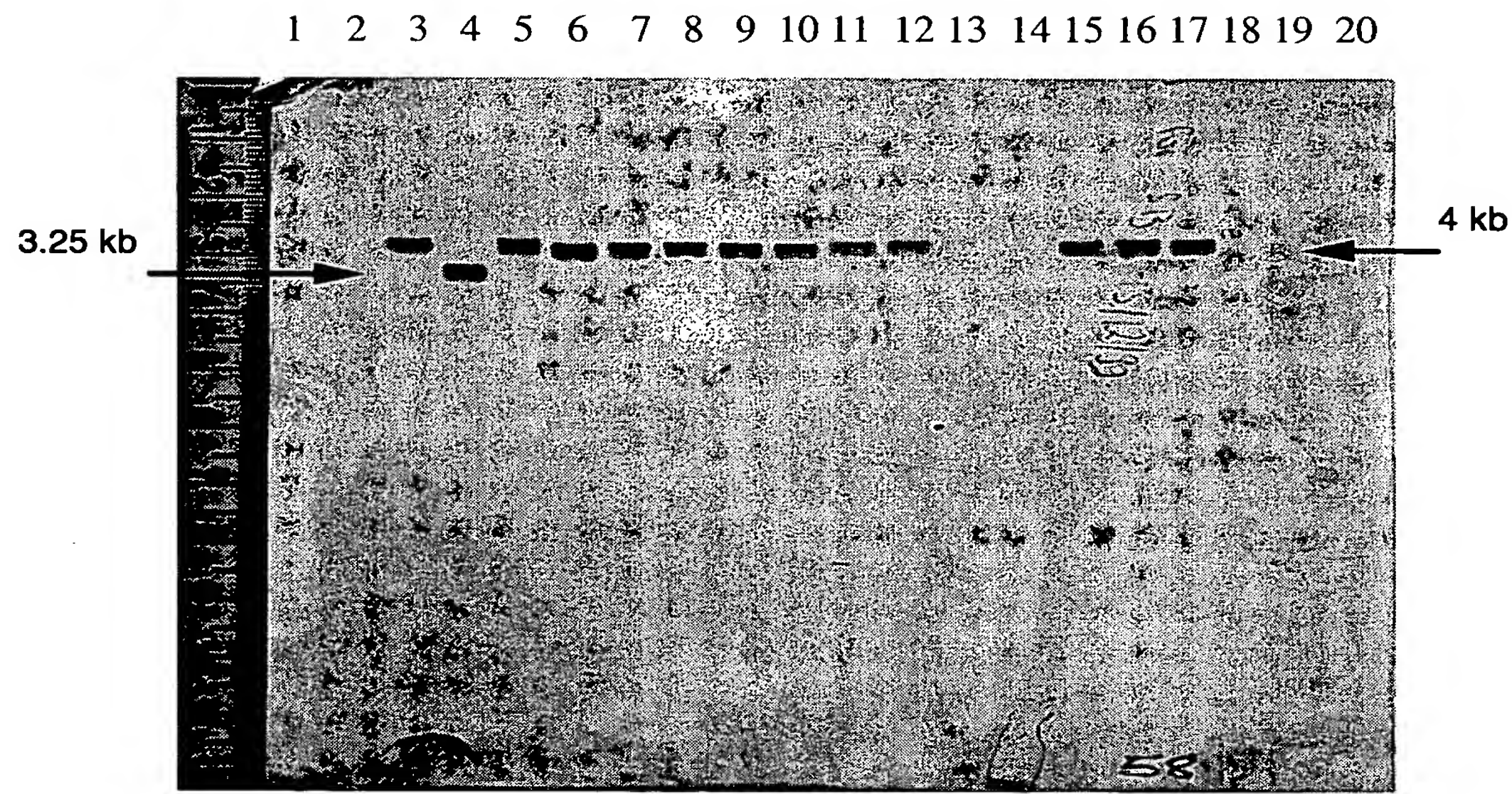
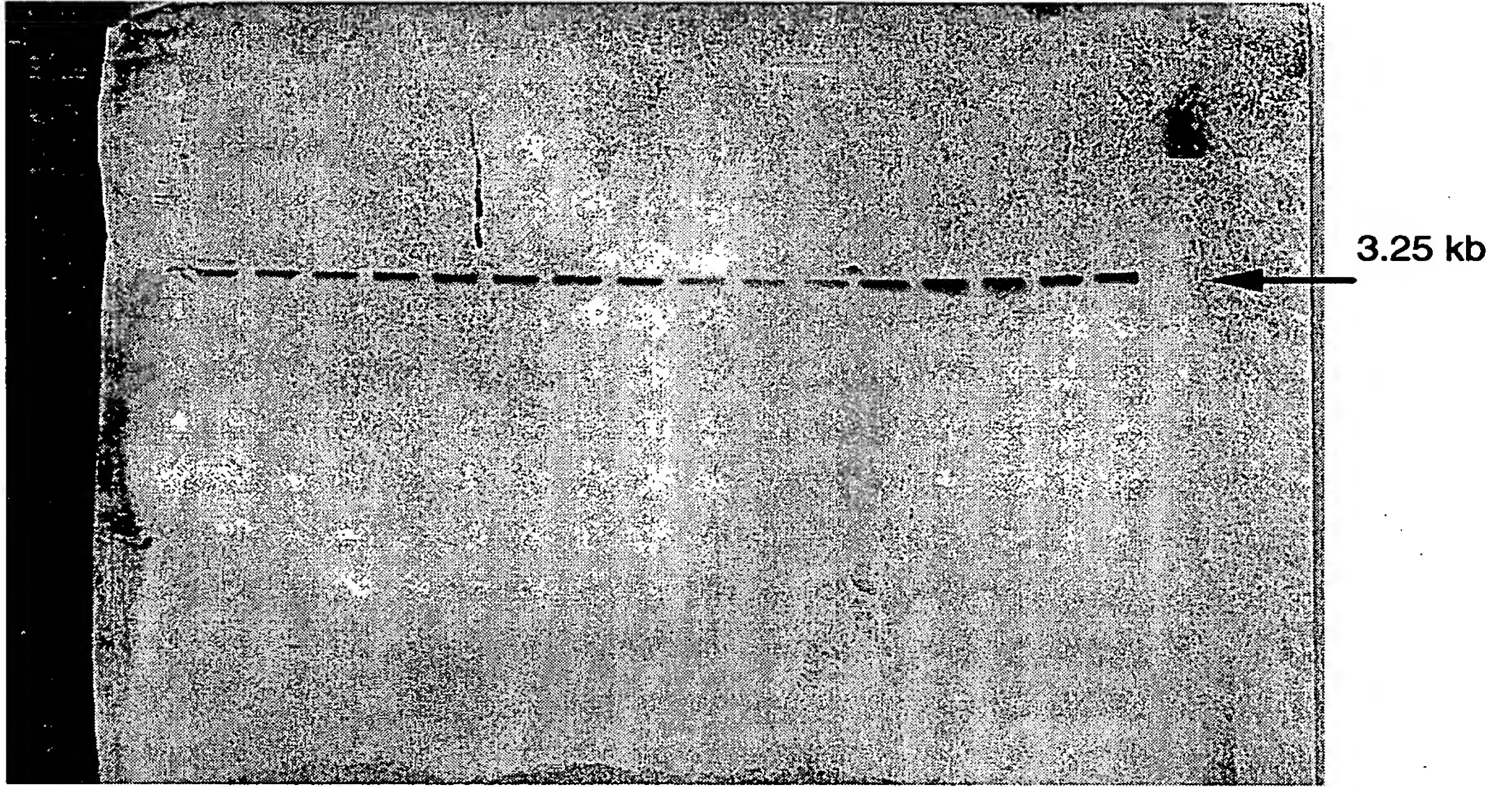




Figure 15 **Southern blot analysis – ID-93**

5 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20



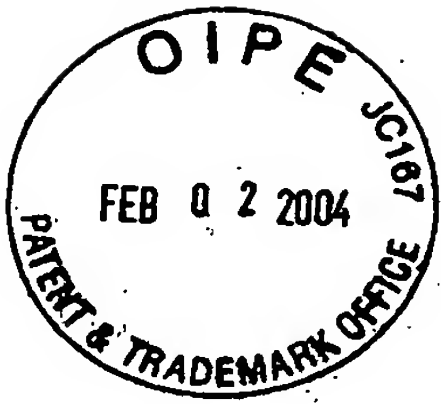


Figure 16 **Southern blot analysis – ID-96**

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

5

